Locating and Sequencing of the E3 and Neighbouring Regions in Bovine Adenovirus Type 2

by

Lesley E. Esford, BSc. (Hons), BAdmin (Hons)
Brock University, St. Catharines, Ontario

A Thesis

submitted to the Department of Biological Sciences

in partial fulfillment of the requirements

for the degree of

Master of Science

July, 1993

Brock University

St. Catharines, Ontario

TABLE OF CONTENTS

Abstract

Adenoviruses are nonenveloped icosahedral shaped particles. The double stranded DNA viral genome is divided into 5 major early transcription units, designated E1A, E1B, and E2 to E4, which are expressed in a regulated manner soon after infection. The gene products of the early region 3 (E3), shown to be nonessential for viral replication *in vitro*, are believed to be involved in counteracting host immunosurveillance. In order to sequence the E3 region of Bovine adenovirus type 2 (BAV2) it was necessary to determine the restriction map for the plasmid pEA48. A physical restriction endonuclease map for *Bam*HI, *Cla*I, *Eco* RI, *Hin*dIII, *Kpn*I, *Pst*I, *Sal*I, and *Xba*I was constructed. The DNA insert in pEA48 was determined to be viral in origin using Southern hybridization. A human adenovirus type 5 recombinant plasmid, containing partial DNA fragments of the two transcription units L4 and L5 that lie just outside the E3, was used to localize this region. The recombinant plasmid pEA was subcloned to facilitate sequencing. The DNA sequences between 74.8 and 90.5 map units containing the E3, the hexon associated protein (pVIII), and the fibre gene were determined. Homology comparison revealed that the genes for the hexon associated pVIII and the fibre protein are conserved. The last 70 amino acids of the BAV2 pVIII were the most conserved, showing a similarity of 87 percent with Ad2 pVIII. A comparison between the predicted amino acid sequences of BAV2 and Ad40, Ad41, Ad2 and Ad5, revealed that they have an identical secondary structure consisting of a tail, a shaft and a knob. The shaft is composed of 22, 15 amino acid motifs, with periodic glycines and hydrophobic residues. The E3 region was found to consist of about 2.3 Kbp and to encode four proteins that were greater than 60 amino acids. However, these four open reading frames did not show significant homology to any other known adenovirus DNA or protein sequence.

Acknowledgments

There are several people that I would like to extend my thanks to: Dr. Y. Haj Ahmad my supervisor for his guidance, financial support and the suggestion of the thesis topic; to my committee members Dr. D. Bruce and Dr. B. Carlone for their suggestions and tremendous support throughout my masters degree; to my examining committee; Dr. M Manocha, Dr. A. Castle, Dr. F. McCarthy, and Dr. M. Brown (University of Toronto) for useful insight and valuable discussion. A special thank you, to Dr. J.C. Lewis for his encouragement and friendship during my years at Brock University.

Nezar Rghei , Bruce Mackay and Mabrouk Elgadi provided useful discussions and help with many techniques. I would like to thank Kirsty Salmon for the gift of the plasmids pEA48 and PEB113 and for work done leading to the construction of the linear map. I would like to extend my thanks to my fellow graduate students also to my family and friends especially my mom and dad who supported and encouraged me always. A special thanks to my mom who fed me and helped me whenever she could, in particular with the rereading of the sequence.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Ad 2, 3, 5, etc.	adenovirus type 2, 3, 5, etc.
BAV 1, 2, 3, etc.	Bovine adenovirus type 1, 2, 3, etc.
bp	base pair
BSA	bovine serum albumin
$CaCl_2$	calcium chloride
Ci	curie
cm	centimeter
CO_2	carbon dioxide
cpe	cytopathic effect
CTL	cytotoxic T lymphocytes
DBP	DNA binding protein
dATP, dCTP, dGTP, dTTP	deoxynucleotide triphosphate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DTT	dithiotheitol
E1 - E4	early region 1 - 4 of adenovirus
EDTA	ethylenediamine tetracetic acid
et al.,	and co-workers
EtBr	ethidium bromide
EtOH	ethanol
FBS	fetal bovine serum
fig.	figure
g	gram
h	hour
HCl	hydrochloric acid
ITR	inverted terminal repeat
kbp	kilobase pair
KCl	potassium chloride
KDa	kilodaltons
KH_2PO_4	potassium phosphate buffer
L1 - L5	late region 1 - 5 of adenovirus genome
L	liter
LB	Luria-Bertani medium
M	molar
MCS	multiple cloning site
MDBK	Madin-Darby Bovine kidney cells
min	minute
$MgCl_2$	magnesium chloride
MEM	minimal essential medium
ml	milliliter
MLP	major late promoter
mm	millimeter
mM	millimolar
MOI	multiplicity of infection
mu	map units
MW	molecular weight

NaAc	sodium acetate
NaCl	sodium chloride
NaOH	sodium hydroxide
$NaPO_4$	sodium phosphate
ng	nanogram
NP - 40	nonidet P-40
OD	optical density
PFU	plaque forming unit
RNA	ribonucleic acid
RNase	ribonuclease
rpm	rotation per minute
RT	room temperature
SDS	sodium dodecyl sulfate
SSC	saline sodium citrate
SV40	simian virus 40
TBE	Tris borate/EDTA buffer
TE	Tris-EDTA buffer
TNF	tumor necrosis factor
Tp	terminal protein
Tris	Tris (hydromethyl) aminomethane
UV	ultraviolet radiation
v/v	volume to volume
^{32}P	phosphorous -32
µg	microgram
µl	microliter

CHAPTER 1: INTRODUCTION

Since adenoviruses were first discovered by Rowe *et al.,* (1953) they have been isolated worldwide from a number of different hosts. The family Adenoviridae is divided into two genera; *Mastadenovirus* and *Aviadenovirus*. All mammalian adenoviruses, including bovine adenoviruses, belong to the genus *Mastadenovirus* . All adenoviruses share a common architecture. The nonenveloped capsid is icosahedral shaped, with long projections called fibres extending from each of the 12 vertices. Only the number and the length of the fibre, vary between serotypes. Also, each virion contains a single molecule of double stranded DNA with inverted terminal repeats and a 55 KDa terminal protein. The organization of the genome also appears to be conserved between serotypes. The viral genome is divided into early and late transcription units. The early regions designated E1 to E4 are transcribed soon after infection. The late regions designated L1 to L4 are transcribed after the onset of viral replication and primarily code for structural proteins.

Adenoviruses are well studied for several reasons. First, adenoviruses were the first DNA viruses to show oncogenic potential. Also, adenoviruses are an excellent model system to investigate cell transformation and eukaryotic gene regulation. In addition, adenoviruses are currently being used as viral vectors in mammalian cells (Haj-Ahmad & Graham, 1986).

The early region 3 (E3) is studied for two reasons. First the E3 gene products are believed to be involved in modulating the host immune responses to

the virus infection (Wold & Gooding, 1991). Secondly, the region is nonessential for viral replication in cultured cells. This has implications for the development of BAV2 based expression vectors for the possible use as vaccines in cattle. Non defective human adenovirus vectors have already been developed by substituting the E3 sequences with foreign DNA (reviewed by Graham & Prevec, 1991).

In this study, the region between 74.8 and 90.5 map units was sequenced and analyzed. Homology comparisons with other adenoviral DNAs indicate that this region contains part of the hexon associated protein (pVIII) precursor on the left and a 59.6 KDa fibre encoding gene on the right of the E3 region. The putative E3 region lies between 76.1 and 83.3 map units and encodes four polypeptides larger than 6 KDa. Although the BAV2 E3 region is smaller than that inhuman adenoviruses, it was larger than the previously sequenced BAV3 E3 (Mittal *et al.*, 1992). The 2.3 kbp E3 region in BAV2 represented 7 percent of the viral genome in contrast to the 1.5 kbp E3 in BAV3 which only represented 3 percent of its viral genome (Mittal *et al.*, 1992).

The following chapter will examine the molecular biology of adenoviruses. Specifically the early region 3 (E3) will be discussed as a prelude to a description of my studies aimed at localizing and characterizing the E3 and neighbouring regions in bovine adenovirus type 2 (BAV2). Much of the precise information, especially in the field of molecular biology, has been accumulated on human adenoviruses (Ad). Therefore, any studies on other mammalian viruses lead to a better understanding of this family of adenoviruses as a whole.

CHAPTER 2: LITERATURE REVIEW

2.1 Discovery and Epidemiology of Adenoviruses

Adenoviruses are a ubiquitous group of viruses that were first isolated by Rowe *et al.,* in 1953 . The first virus was isolated from human adenoid tissue hence the name adenovirus (Enders *et al.,* 1956). Since 1953 over 80 serotypes have been identified of which 47 are of human origin (Roos, 1989). Adenoviruses can cause both mild respiratory and gastrointestinal infections in humans (reviewed by Horwitz, 1990 a, b). Also, several adenoviruses (Ad) are oncogenic for a number of newborn rodents (reviewed by Horwitz, 1990 a). The demonstration of the first oncogenic DNA virus, as well as other unique observations first made on adenoviruses, have made this group a popular research tool in molecular biology. Adenoviruses have been used by many workers to investigate such phenomena as cell transformation and eukaryotic gene regulation (Zock & Doerfler, 1990). In both of these areas adenoviruses have increased our understanding of the molecular mechanisms involved in these processes. In addition, adenoviruses are currently being used as viral vectors in mammalian cells (Haj - Ahmad & Graham, 1986; Venkatesh *et al.,* 1990).

2.2 Classification

The familyAdenoviridae is divided into two genera; *Mastadenovirus* for the mammalian adenoviruses and *Aviadenovirus* for the avian viruses. Members of

each genus share a group specific antigen, although there is no common antigen that characterizes the whole family (Horwitz, 1990 a). The 47 human serotypes are grouped into six subgroups (A to G) on the basis of antigenic cross reactivity and other criteria (Wold & Gooding,1991).

2.3 Bovine Adenoviruses

Since bovine adenoviruses (BAV) were first discovered (Klein *et al.,* 1959), they have been isolated world wide from cattle, sheep (Belak & Palfi, 1974, a, b), free living buffalo (Baber & Condy, 1981), fallow deer (Boros *et al.,* 1985) as well as apparently healthy cattle (Darbyshire *et al.,* 1966 b; Rondhuis. 1968; Burki *et al.,* 1978). The study of BAV is of economical importance because BAV infections can result in poor weight gain and reduced milk production. Like other adenoviruses, BAV infections have been associated with respiratory tract disease, enteritis, conjunctivitis, keratitis and pyrexia (Mohanty, 1971; Mattson, 1973), that could result in weak calf syndrome in newborn calves and adult cattle (Lehmkuhl *et al.,* 1975; Key & Derbyshire, 1984; Scarziani *et al.,* 1989).

The ten serotypes of BAV are divided into two groups, based on serum neutralization and complement fixation tests as well as the type of cell culture required for replication (Bartha 1969; Horner *et al.,* 1989). Subgroup I, which includes serotypes 1, 2, 3, and 9, carries the antigen common to mastadenoviruses and grows in bovine kidney cells as efficiently as in bovine testicle cells (reviewed by Ishibashi & Yasue, 1984). In contrast subgroup II which includes serotype 4, 5, 6, 7, 8 and 10 carries so little of the common antigen that it was hardly detectable and could only be propagated in bovine testicle cell cultures (reviewed by Ishibashi & Yasue, 1984). However, recently

Benko *et al.* (1989) found that 9 out of 10 BAV serotypes would grow efficiently in calf thyroid cell culture.

The oncogenic potential of BAV3 has made it the most widely studied of the BAV serotypes. Of the 10 serotypes identified, only BAV3 causes oncogenic transformation in newborn hamsters and mouse cell lines (Darbyshire,1966 a; Igarashi *et al.,* 1978; Motoi & Ogawa, 1985). Like human adenoviruses studied, the transforming region in BAV3 has been localized to the left end of the viral genome between 0 and 11.9 map units (Igarashi *et al.,* 1978; Niiyama *et al.,* 1981). This is not surprising, because hybridization studies carried out between BAV3 and Ad2 revealed extensive DNA homology between the two viruses (Hu *et al.,* 1984). This suggests that the two viruses share similiar genomic organization. Interestingly, BAV, like human Ad, show the same inverse relationship between guanine and cytosine (C+G) content and oncogenic potential. The 48 % C+G content of BAV3 is low in comparison to the nononcogenic BAV1 and BAV2, with 62 and 61 % C+G content, respectively (Panigraphy *et al.,* 1977).

BAV2, first isolated by Klein *et al.* (1960) was used in this study. Two subspecies of BAV2 have been proposed based on hemagglutination properties and restriction enzyme cleavage patterns. Subspecies A refers to isolates similiar to the prototype (#19), whereas subspecies B are more closely related to ORT-111, a strain of BAV2 originally isolated from sheep (Belak *et al.,* 1983). The main difference between these two subspecies lies within the region that is equivalent to the E3 region in Ad 2 (Belak *et al.,* 1986). The size of the genome reported varies from 30.3 to 34.8 kbp (Belak *et al.,* 1983; Belak *et al.,* 1986; *et al.*, 1988) and restriction maps of BAV2 (#19) of *Bam*HI and *Eco*RI are available (Belak *et al.,* 1986).

As previously mentioned most molecular studies have been on human adenoviruses. Therefore, studies non human members of the family must be conducted, not only for the advancement of veterinary medicine but also for a thorough understanding of the family of adenoviruses as a whole.

2.4 Structure of the Virus

All adenoviruses have a common architecture(fig. 1). The virion is composed of 252 capsomers arranged in icosahedral symmetry that form a capsid of 65- 80 nm in diameter (Horne *et al.,* 1959). Two hundred and forty of the capsomers are called hexons because each is surrounded by 6 capsomers. The remaining 12, called pentons, are situated at the 12, 5 fold vertices. The pentons are comprised of a penton base anchored in the capsid and a non-covalently attached fibre that appears as a long projection extending from the vertices. The length of the fibre is characteristic of a given adenovirus serotype and varies from approximately 10nm for Ad3 to almost 50 nm for some avian adenoviruses (Norrby, 1969). Almost all of the adenovirus serotypes studied have a single fibre, however, some types of fowl adenoviruses (Norrby, 1969) and Ad40 and Ad41(Kidd *et al.,* 1993) were found to have two fibers, one short and one long emanating from each penton base. The hexon, penton and fibre, contain the antigenic determinants that are important in the serological classification of adenoviruses . The hexons, pentons and fibre are derived from different viral polypeptides. The viral capsid is comprised of between 11- 15 polypeptides (fig. 1).

2.5 Genome Structure and Organization

Each virion contains one linear double stranded DNA molecule ranging in size from 30-36 kbp, depending on the serotype (Tooze, 1980). The genome of

the adenovirus is divided into left and right halves based upon its G + C content. The strand transcribed to the right is called the r-strand, whereas the strand transcribed to the left is termed the l-strand.

The genome is further subdivided into 100 map units from left to right. Figure 2 demonstrates the principle transcriptional organization of the adenovirus type 2 (Ad2) genome, which is generally similar for the different adenovirus serotypes. The viral genome is divided into early (E) transcription regions and late (L) transcription regions. Early and late refer to the time the region is transcribed relative to the replication of the viral genome. The late regions are transcribed after replication starts and they mainly code for structural proteins.

Two novel features of the viral DNA are: the inverted terminal nucleotide sequence repeats (ITRs) consisting of between 52-165 base pairs on each strand and a 55 KDa protein covalently attached to each 5' end of the DNA molecule. Both the ITR and the 55 KDa protein are believed to play an important role in viral replication (Challberg and Kelly, 1979).

2.6 The Lytic Cycle

Most studies on the lytic cycle have been carried out in Ad2 and Ad5 infected human cell lines. However, the essential features of multiplication are similiar for all serotypes.

The productive cycle is initiated approximately one half hour following infection, when the viral DNA reaches the infected cell nucleus. Sequentially regulated transcription, from early regions E1a, E1b and E3, located on the r-strand and E2a, E2b and E4 on the l -strand, begins prior to viral replication (fig. 2). These early proteins carry out a variety of functions that prepare the cell for efficient virus replication. All viral genes are transcribed by cellular RNA polymerase II, except for the virus associated (VA) RNA I and VA-RNA II which

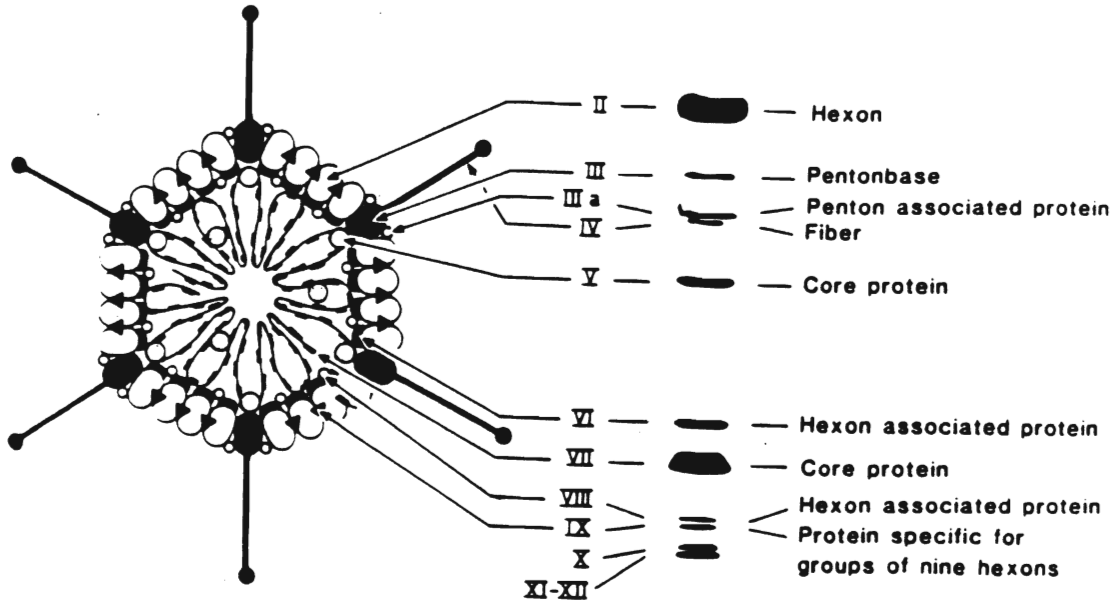


Figure 1:

A schematic representation of an adenovirus. The position of each of the polypeptides is indicated . The hexon (II), penton base (III), fibre (IV) and hexon associated proteins (IIIa, VI, VIII, and IX) are subunits of the capsid. The core contains proteins V, VII, and u, as well as the 55 KDa terminal protein covalently linked at each of the 5' ends of the linear DNA (Reproduced from Horwitz, 1990 a).

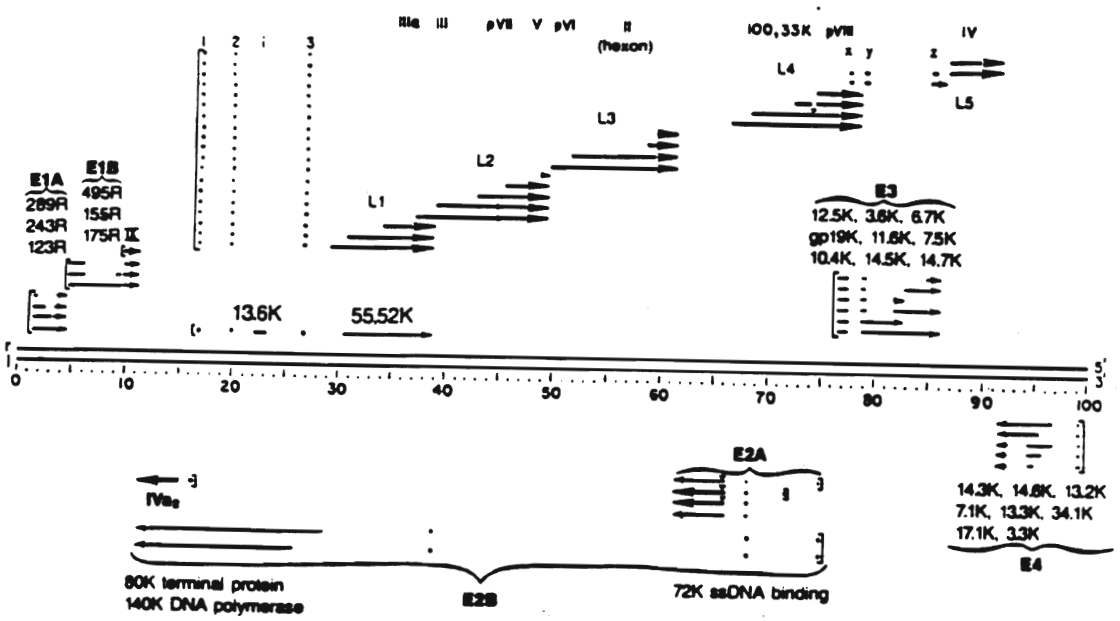


Figure 2:

A schematic drawing showing the principle transcriptional organization of Ad2 genome. The early mRNAs are designated E and are shown as thin lines. The late mRNAs are designated L with heavier lines. The polypeptides are denoted using their molecular weight or by roman numerals (reproduced from Wold and Gooding, 1991)

are transcribed by RNA polymerase III (Breathnach & Chambon, 1991).

Each of the primary transcripts gives rise to a family of mRNA following a RNA splicing mechanism, first detected in the adenovirus system, and found in all higher eukaryotes (Breathnach & Chambon, 1991). All of these early mRNAs are translated to give rise to early proteins , with the sole exception of the VA-RNA and the L1 transcript.

Viral DNA replication begins about 7 hours post infection and this initiates the late phase. Most late genes fall into five families, L1 to L5, coded within the major late transcription unit. Late genes encode mainly virion structural proteins. Host DNA, mRNA and protein synthesis are shut off in late stages of infection. The replication of viral DNA proceeds by a semi-conservative, strand displacement mechanism initiating at or near either termini of the genome. As was previously mentioned, it appears that both the 55 KDa terminal protein (Tp) and the ITR are involved in the replication process. It has been suggested that the 55 KDa Tp is involved in the initiation and elongation process and the down regulation of its own, as well as other early gene expression, presumably by decreasing the stability of the mRNA (Babich & Nevin, 1981)

At about 24 hours post infection, virions assemble in the infected cell nucleus. Approximately 10^5 - 10^6 progeny viral genomes are made, of which only about 20 percent are packaged into viral particles (reviewed by Horwitz, 1990a). The assembly of the virion commences at about 12 to 15 hours post infection in the nucleus of the infected cell.

One of the fundamental structures of the virions is the capsid which encloses the nucleic acid. The capsids are rarely much greater than the volume of genome they can hold. The tolerance or limit for packaging greater amounts of nucleic acid other than the viral genome size is therefore limited to the size of the isometric capsid. Generally, it has been observed that isometric capsids can only

package about 5 percent over the normal genomic size (Haj-Ahmad, 1986). This packaging constraint, as well as the fact that the E3 can be deleted to increase the amount of foreign DNA that can be inserted (Haj-Ahmad & Graham, 1986), are important factors in the use of adenoviruses as viral vectors.

2.7 The Early Region 3 (E3)

The early proteins of adenoviruses, which are initially synthesized prior to viral DNA replication, are of great interest because they exert their functions on key regulatory aspects of the virus and host. Human adenoviruses can sustain and establish latent infections (Fox *et al.,* 1977). Thus, the mechanisms evolved to evade the host response are of great interest. The early region 3 (E3) region is a cassette of genes that down regulates the host immune response. Recently, considerable information has been obtained about the E3 region. The E3 region has been sequenced in Ad2 (Herisse *et al.,* 1980; Herisse & Galibert, 1981), Ad5 (Cladaras & Wold, 1985), Ad3 (Signas *et al.,* 1986), Ad11 (Mei & Wadell, 1992), mouse adenovirus type 1 (MAV1) (Raviprakash *et al.,* 1989; Beard *et al.,* 1990), canine adenovirus type 1 (CAV1) (Dragulev *et al.,* 1991; Linne, 1992) and CAV2 (Linne, 1992), bovine adenovirus type 3 (BAV3) (Mittal *et al.,* 1992) and partially in Ad7 (Hong *et al.,* 1988), and Ad35 (Flomenberg *et al.,* 1988). The human E3 region is located between 76.6 and 86 map units on the r-strand (fig. 3), and encodes at least nine mRNA species (Chow *et al.,* 1979; Cladaras *et al.,* 1985; Cladaras & Wold, 1985; Herisse *et al.,* 1980). These Ad2/Ad5 mRNAs, as well as the predicted protein products, are depicted in figure 3. Six of these proteins have been identified and functions are known for four of them. These are 6.7 KDa (Wilson-Rawls *et al.,* 1990), gp19 KDa (also called E19); (Persson *et al.*,1980),11.6 KDa (Wold *et al.,* 1984), 10.4 KDa (Tollefson *et al.,* 1990a), 14.5 KDa (Tollefson *et al.,* 1990b) and 14.7 KDa (Tollefson and Wold, 1988, Wang *et*

al., 1988). All of these proteins except perhaps the 6.7 KDa are found in group B adenoviruses (Wold & Gooding, 1991). The non human adenoviruses studied did not show any substantial amino acid or DNA sequence homology to the human Ad E3 proteins (Raviprakash *et al.*, 1989; Beard *et al.*, 1990; Dragulev *et al.*, 1991; Linne, 1992; Mittal *et al.*, 1992).

All of the E3 messages share a common 5' leader, whereas the 3' ends of these transcripts vary (Cladaras & Wold, 1985; Cladaras *et al.,* 1985). The E3 region of MAV1, though significantly smaller than Ad5, shows a similiar splicing pattern. The transcription mapping of the other non human adenoviruses has not been determined.

The first major E3 protein identified and characterized was the transmembrane glycoprotein (gp19 KDa), localized in the endoplasmic reticulum. Specifically, the gp19 KDa glycoprotein encoded by the E3 binds strongly with newly synthesized class I antigens of the major histocompatibility complex and impedes their transport to the cell surface (Andersson *et al.,* 1985; Burgert & Kvist, 1985). The formation of the complex between E3 gp19 KDa and newly formed class 1 antigens prevents their terminal glycosylation and transport to the cell surface (Andersson *et al.,* 1985; Burgert & Kvist, 1985). The reduced surface levels of class I antigens could prevent the infected cell from being recognized by cytotoxic T lymphocytes and increase the potential for establishing latent infection (Andersson *et al.,* 1985; Burgert & Kvist, 1985; Burgert *et al.,* 1987). Consistent with this, in cells expressing gp19 KDa, is the lowered expression of class I antigens correlated with decreased lysis by cytotoxic T cells (CTL) specific for class I antigens (Burgert *et al.,* 1987). All adenoviruses can modulate cell-surface expression of class 1 antigens (Paabo *et al.,* 1986) and the inhibition of CTL target cell lysis maps exclusively to the gp19 KDa gene in Ad5 (Rawle *et al.,* 1989). Furthermore, it has been suggested that gp19 KDa may be

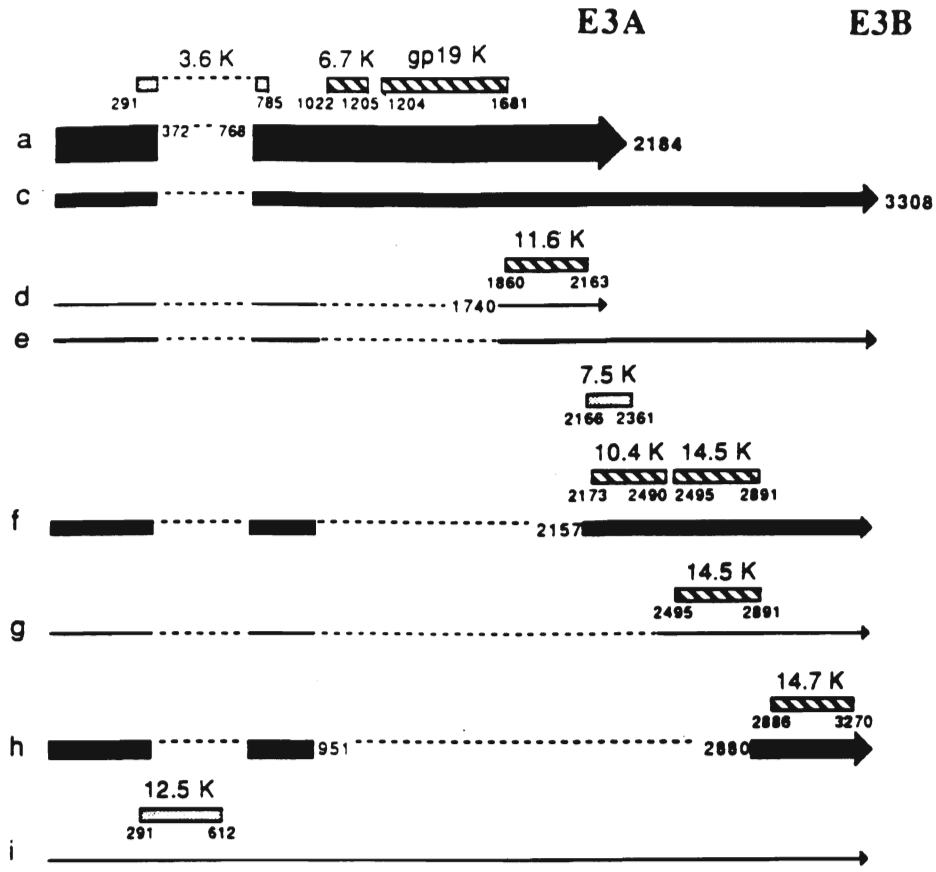


Figure 3:

A schematic illustration of the E3 region of Ad2. Arrows indicate the spliced structures of the mRNAs; the thickness of the arrow reflects the relative abundance of the mRNAs and the dashed lines indicate introns. E3A and E3 B are polyadenylation sites. Bars above the arrows indicate proteins; hatched bars are proteins that have been identified, and stippled bars are proteins proposed to exist (reproduced from Wold & Gooding, 1991).

involved in suppressing the production of cytokines which attract lymphocytes, because there is an increase in the extent of the lymphocyte and macrophage/monocyte inflammatory response in the lungs of cotton rats infected with Ads mutants that lack the gp19 KDa (Ginsberg *et al.,* 1989).

In addition, recent studies on Ad5 suggest that a difference in E1a protein levels might account for the difference in the ability of CTL to lyse Ad infected cells. Zhang *et al.* (1991) found that E1a protein products were expressed at substantially lower levels during wild type Ad5 infection compared with E3 defective mutants. They concluded that E3 products interfere with the ability of to lyse the cell indirectly by reducing E1a antigen expression via some post transcription modification.

Although CTL are believed to be a primary means by which the host counteracts virus infections, tumour necrosis factor (TNF) appears to be another. TNF is a multi-functional protein hormone that is secreted by macrophages and monocytes in response to inflammatory stimuli, that has antiviral properties *in vitro* and that may counteract virus replication *in vivo* (reviewed by Horton *et al.,* 1991). Ad-infected mouse lines are sensitized to TNF cytolysis by expression of the proteins encoded by the E1a transcription unit (Duerksen- Hughes *et al.,* 1989; 1991). Recent studies have suggested that 14.7 KDa is expressed during infection (Tollefson & Wold, 1988; Wang *et al.,* 1988) and is necessary to prevent TNF cytolysis in adenovirus infected cells (Gooding *et al.*, 1988). This anti-TNF function is manifested by serotypes A, B, C, D and E, indicating that this is a general property of Ads in these groups (Horton *et al.,* 1990). Also, Horton *et al.* (1991) found that 14.7 KDa, in the absence of other adenovirus proteins can protect mouse cell lines against TNF cytolysis. The mechanism of the14.7 KDa suppression of TNF cytolysis remains unclear although, Horton *et al.* (1991) suggest that it acts as an intracellular pathway specific for the TNF induced lytic

process. It has been suggested that the apparently ubiquitous 14.7 KDa protein, found to be conserved among the serotypes like the gp19 KDa, is another mechanism used by adenovirus to avoid the immune system *in vivo*.

Gooding *et al.* (1991 b) found that in the absence of the 14.7 KDa protein, two other proteins, the 10.4 KDa and the 14.5 KDa, can also protect many mouse cell lines against TNF. These two proteins function together to preventthe lytic activity of TNF. Tollefson *et al.* (1991) concluded that these two proteins form a stable complex because the 10.4 KDa efficiently coimmunoprecipitates with the 14.5 KDa and both are translated from mRNA f, which suggests a functional relationship. The 10.4/14.5 KDa complex was found to suppress TNF killing in 11 of the 15 mouse cell lines tested . Both proteins are cytoplasmic integral proteins as predicted from their sequence (Cladaras & Wold , 1985) and as shown experimentally (Tollefson *et al.,* 1990 b; Krajcsi *et al.,* 1992 a, b; Hoffman *et al.,* 1992). Both proteins are oriented in the membrane with their C-termini extending into the cytoplasm (Krajcsi *et al.,* 1992 a, b). This C terminal region of 14.5 KDa is the most conserved among Ad5, Ad2, Ad3 and Ad7 and it is essential for 14.5 KDa to prevent TNF cytolysis (Krajcsi & Wold, 1992 c).

Human cells are even more complicated in their response to adenovirus infection and TNF. The 19 KDa protein derived from the E1b transcription unit is a third protein that can prevent TNF cytolysis; E1b- 19 KDa provides protection against TNF in Ad-infected human cells but not in mouse cells (Gooding *et al.,* 1991a).

Considering that three sets of proteins (the E3 14.7 KDa and 10.4 KDa /14.5 KDa proteins and the E1b 19 KDa protein) protect Ad infected cells against TNF cytolysis it would appear that TNF is a major anti-adenovirus

defense of the host (Gooding *et al.,* 1988; 1991a, b). It is likely that the 14.7 KDa and the 10.4 KDa /14.5 KDa protect human cells in *vivo* .

The 10.4 KDa protein was first associated with the down regulation of the epidermal growth factor receptor (EGF-R) (Carlin *et al.,* 1989). The EGF receptor is a transmembrane glycoprotein and it is a member of the tyrosine kinase class of membrane receptors, localized on the plasma membrane (reviewed by Tollefson *et al.,* 1991). EGF is a polypeptide growth factor. The formation of the EGF/EGF-R complex and its subsequent action activates cellular metabolism and eventually induces DNA synthesis and mitosis (reviewed by Tollefson *et al.,* 1991). The tyrosine kinase activity of EGF-R is essential for EGF signal. The 10.4 KDa and the 14.5 KDa protein were first characterized by Tollefson *et al.,* (1990a, b). Both proteins are required for the down regulation of the EGF-R (Carlin *et al.,* 1989; Tollefson *et al.,* 1991). Down regulation in this context means that the cell surface EGF-R was internalized in a mediated pathway and degraded, presumably in lysosomes. However, it is unclear what mechanisms are used and the role that both proteins play. This additional function further supports that there is a relationship between these two proteins. It is not known whether the effects of 10.4/14.7 on EGF-R are related to the effects on TNF cytolysis (Krajcsi *et al.,* 1992 a)

In addition to the proteins previously mentioned, the 11.6 KDa (Wold *et al.,* 1984), 6.7 KDa (Wilson-Rawles *et al.,* 1990) and the 12.5 KDa proteins (Hawkins & Wold, 1992) have also been shown to exist. The 11.6 KDa protein is unique among the E3 proteins in that it is made in low amounts during early stages but very abundantly at late stages, when its mRNA is formed by splicing from the pre-mRNA of the major late transcription unit (Tollefson *et al.,* 1992). This would suggest that the 11.6 KDa must not only be a member of the E3 region but also of the late transcription unit. In contrast to the 12.5 KDa and the

6.7 KDa proteins that are well conserved in whole or in part in Ad2 and Ad5 (Group C) and Ad3 (group B; Wilson-Rawls *et al.,* 1990; Hawkins & Wold, 1992) the sequence of the11.6 KDa protein is more divergent. The functions of the 11.6 KDa, 12.5 KDa and 6.7 KDa proteins are presently unknown.

There are two main reasons why the E3 region is studied. The first can be inferred from the previous discussion; characterizing the E3 region may lead to a better understanding of the host virus interaction. Secondly, the E3 region has been shown to be non-essential for the Ad growth in tissue culture or for viral replication *in vivo* (Haj-Ahmad & Graham, 1986; Morin *et al.,* 1987; Ginsberg *et al.,* 1989). Deletions in this region appear to have no effect on viral replication. This findings have two implications. First, although this region can be deleted, it is always found in natural isolates of adenoviruses (Adrian *et al.,* 1989 a, b) suggesting that these genes are important for natural infections of the host. The E3 region has probably been maintained in numerous serotypes because of its function in modulating the host response to Ad infection (Ginsberg *et al.*, 1989). Secondly, if BAV2 has an E3 region and it can be deleted this will led to the development of a viral vector. The possibility of substituting E3 sequences with foreign DNA has led to the development of nondefective Ad vectors which can replicate and express foreign genes in cultured cells and in infected animals (reviewed by Graham & Prevec, 1991) .

The E3 region is one of the least conserved regions among adenovirus serotypes. The two regions that flank the E3 region, the L4(on the left) and the L5 (on the right), will be critical in locating the E3 region in BAV2.

2.8 The Late Region 4 (L4)

The L4 region encodes for three proteins, 100 KDa , 33 KDa and protein VIII. The 100 KDa, and the 33 KDa are non-structural proteins, where as protein pVIII

is a structural protein. The DNA sequence and the amino acid sequence of pVIII have been shown to be highly conserved amongst Ad serotypes (Raviprakash *et al.*, 1989; Dragulev *et al.*, 1991; Chroboczek *et al.*, 1992; Mittal *et al.*, 1992). It is not surprising that this region is conserved because pVIII is an internal capsid protein. However, not much is known about the structural and functional aspects of pVIII.

2. 9 The Late Region 5 (L5)

The L5 encodes a single protein called the fibre. The fibre is responsible for attachment of the adenovirus to the cell surface (Phillipson *et al.,* 1968) and contains a large fraction of the antigenic sites of the virion (reviewed in Phillipson, 1983). The fibre is made up of two (Green *et al.,* 1983) or three (Dorsett & Ginsberg, 1975; Stouten *et al.,* 1992) identical polypeptides. Analysis of the primary structures of fibers from several adenovirus serotypes [Ad2 (Herisse *et al.,* 1981), Ad3 (Signas *et al.,* 1985), Ad5 (Chroboczek & Jacrot, 1987), Ad7 (Hong *et al.,* 1988), Ad40 (Kidd & Erasmus, 1989), Ad41 (Pieniazek *et al.,* 1989) Ad11 (Mei & Wadell, 1993), MAV1 (Raviprakash *et al.,* 1989) and CAV1 (Dragulev *et al.,* 1991)] and other experimental data, suggest an almost identical overall structure of the fiber. It consists of a tail, which associates in a noncovalent manner with the penton base (Devaux *et al.,* 1987; Weber *et al.,* 1989), a shaft of variable length depending on the serotype (Norrby, 1969) and a distal knob.

Green *et al.* (1983) proposed a model for the organization of the shaft based on the observation of a periodic repeat of hydrophobic residues between amino acids 40 and 400 in Ad2. Each motif of 15 residues is organized in two short β-sheets and two β-bends, the entire structure forming a long narrow amphipathic sheet. The number of repeating motifs in the shaft region varies

between Ad serotypes from 6 to 42 (Signas *et al.,* 1985; Chroboczek & Jacrot, 1987; Hong *et al.,* 1988; Raviprakash *et al.,* 1989; Kidd *et al.,* 1990; Dragulev *et al.,* 1991; Mittal *et al.,* 1992; Kidd *et al.,* 1993; Mei & Wadell, 1993). The fibre sequences of several adenoviruses support this model .

From sequence comparisons (Chroboczek & Jacrot, 1987) the most conserved region within the fiber is the amino terminal (tail) portion. It is not surprising that this region is completely conserved between Ad2 and Ad5 because of evolutionary constraints on the interaction between fibre tail and the penton base. In contrast the shaft and knob are exposed and the nucleotide and amino acid sequences are variable even among serotypes that belong to the same group. It is not expected that BAV2 will show significant nucleotide identity to any fibre sequence known.

2.10 Viral Base Vectors

One of the more recent and most efficient methods for the reintroduction of altered DNA using recombinant DNA technology has been through the use of viral based vectors. Among the advantages of viral vectors is the high efficiency of gene transfer *in vitro* and possibly *in vivo*. Some of the possible uses of mammalian viral vectors are: the study of gene expression, the correction of a genetic disease through gene therapy, or the genetic engineering of vaccines . Several viruses have been used, including retroviruses, poxviruses, herpes viruses, and polioviruses (reviewed by Panicali *et al.,* 1983; Glutzman, 1984). However, the use of adenoviruses as viral vectors has become increasingly popular. The adenovirus has been used as a vector for the expression of many viral proteins, including HBsAg, HIV envelope proteins, herpes virus proteins and vesicular stomatitis virus proteins (reviewed in Ye *et al.,* 1991).

Adenoviruses are particularly well suited as gene transfer vectors in mammalian cells for several reasons. First, they are widely studied and well characterized both genetically and biochemically; secondly, they are easy and inexpensive to grow and manipulate and exhibit a broad host range *in vitro* and *in vivo*; third, copious amounts of virus and viral products can be produced in lytically infected cells (Haj-Ahmad & Graham, 1986). In addition, since only a small portion of the viral genome appears to be required in *cis*, adenovirus derived vectors offer excellent potential for the substitution of large fragments once cell lines, analogous to the 293, have been developed which can provide most of the essential viral functions in *trans* (Haj-Ahmad & Graham, 1986). Finally, most adenovirus infections are relatively mild. However, the safety of adenoviruses in the development of novel vaccines has recently been questioned because of the ability of Ad to cause transformation in cell lines and to cause tumours. Current vaccine strains, Ad4 and Ad7, have no oncogenic potential for rodents. Some of these factors represent unique advantages adenoviruses have over vectors derived from retroviruses and other DNA viruses.

As mentioned previously, due to packaging constraints, only 2 kbp of foreign DNA can be inserted into the viral genome. The removal of the E3 region can increase this limitation to approximately 5 kbp (Haj-Ahmad & Graham, 1986). The regions of the Ad genome that have been used for the insertion of foreign genes are the E1 and the E3 regions. If the E1 region is deleted, the recombinant genome must be propagated as a homogenous virus by infecting the 293 cell line (Haj-Ahmad & Graham, 1986). The special HEK-293 cell line contains the complements of the Ad E1 region. These Ad vectors would be defective in cells other than the 293 line. However, as the E3 region can be deleted and replaced without compromising the viability of the virus in tissue culture, an associated helper virus is not required. The successful insertion and

expression of the hepatitis B surface antigen in the Ad E3 region with subsequent inoculation and antibody formation in the hamster has been reported recently (reviewed by Horwitz, 1990 a).

The application of adenoviruses as viral based vectors is in its formative years. The use of viral vectors to study gene expression is critical. However, the use of Ad viral based vectors in gene therapy and recombinant viral vaccines will be slow. Although this virus family is relatively well characterized much of the molecular biology of this virus is presently being discovered. For example, most of the information about the E3 region has been collected over the last three years. Prior to 1988 only the E3 19 KDa protein was well characterized.

2.11 Purpose of this Investigation

The three objectives of this study are: first, to construct a detailed restriction endonuclease map of the recombinant plasmid pEA48 which contains the BAV2 *Eco*RI A fragment; secondly, to localize the E3 region within the BAV2 genome using Southern hybridization techniques and thirdly, to sequence and partially analyze the E3 and neighboring regions.

CHAPTER 3: METHODS AND MATERIALS

3.1 Cell Line and Viral Strain

Madin-Darby bovine kidney (MDBK) cells (a generous gift from J.B. Derbyshire, University of Guelph , Guelph, Ontario, Canada) were grown and maintained in 150 mm plates in alpha-minimum essential medium (α-MEM) supplemented with 10% fetal bovine serum (FBS), L-glutamine (292.0 mg/L), 3% sodium bicarbonate and 1.5% antibiotic-antimycotic (Gibco Laboratories). The cells were grown in a humid 37°C incubator under constant atmospheric pressure of 5% CO_2 until monolayers were visible. To maintain the cell line the cells were passaged once a monolayer was visible. The medium was aspirated and the cells were rinsed twice with 1-2 ml of warm versene (0.68 mM EDTA, 0.14 M NaCl, 2.68 mM KCl, 0.81 mM Na_2HP0_4, 1.11 mM glucose, and 1.47 mM KH_2P0_4). The plates were incubated at 37°C with 1 ml of versene until the cells became suspended (approximately 15 min). The cells were collected with a sterile glass pipette, added to an appropriate amount of supplemented α-MEM, then allocated into new plates.

Bovine adenovirus type 2 (BAV2) strain #19, purchased from the American Type Culture Collection (ATCC), was used in this study.

3.1.1 Propagation of BAV2

One hundred and fifty mm dishes of MDBK cells were infected with approximately 10 plaque forming units (pfu) /cell of BAV2 and maintained in α-MEM with 5% FBS. Infected cells were left at 37°C in the humid CO_2 incubator until 80-90% cytopathic effect (CPE) was observed at which time the cells were harvested according to the method described by Hirt (1967).

The plates were scraped with a rubber policeman, the suspension was placed in a sterile 50 ml plastic Corning tube and was centrifuged at 800 rpm (IEC Centra-7R) for 15 min at 4°C. The cell pellet was treated with 0.45 ml of lysing buffer (0.01 M Tris; 0.01 M EDTA; 0.4% sodium dodecyl sulphate (SDS); pH 8.0) and 0.5 ml of a 10X pronase stock (5 mg/ml in 0.01 M Tris; pH 8.0) and incubated at 37°C for 8 to 12 hours, followed by treatment with equal volumes of phenol:chloroform:isoamyl alcohol (25:24:1 v/v). The DNA was recovered by adding 2 volumes of cold ethanol to the aqueous phase, incubating at -20°C for 15 min and centrifuging at 12,000 rpm for 5 min at room temperature. The ethanol was aspirated and the DNA was resuspended in an appropriate volume of TE buffer (10 mM Tris.HCl, 1 mM EDTA, pH 8.0) and stored at -20°C.

3.2 Construction of Recombinant Plasmids

The plasmids pEA48 (containing the viral *Eco*RI A fragment) and pEB113 (containing the *Eco*RI B fragment) were a kind gift from Kirsty Salmon (Brock University, St. Catharines, Ontario, Canada). Other recombinant plasmids were generated to facilitate sequencing by digesting pEA48 DNA and pUC19 DNA with the same restriction endonuclease. Digested plasmid DNA (10-50 ng) was mixed with 100 ng of linearized pUC19 in the presence of ligation buffer (10 mM MgCl2; 20 mM dithiothreitol; 50 mM Tris; 1 mM ATP; pH 7.5, supplied by NEB or Promega), 1-3 units of T4 ligase (NEB or Promega) and made up to volume

(typically 30 µl) with distilled water. A typical reaction was carried out for 1-2 hours at room temperature or overnight at 16°C. The ligation mixture was stored at -20 °C until it was used in the transformation of a strain of *Escherichia coli* (*E. coli.*).

3.3 Bacterial Strains

Escherichia coli strain DH5α (Bethesda Research Laboratories, BRL) was used as a host for most of the constructed recombinant plasmids. Its genotype is *supE*44 Δ*lacU*169 (ϕ80 *lacZ* ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1. The methylation deficient *E. coli* strain GM48 (*dam*$^-$/*dcm*$^-$) (a gift from F.L. Graham, McMaster University, Hamilton, Ontario, Canada), was also used as a host for some of the recombinant plasmids. Its genotype is *thr leu thi lacY galK galT ara tonA tsx dam dcm supE*44.

3.3.1 Propagation and Maintenance of Bacterial Cultures

All bacteria were grown in sterile Luria broth (LB; 10% Bacto-Tryptone, Difco; 5% Bacto-Yeast Extract, Difco; 10% NaCl, Sigma, pH 7) with continuous shaking in either a floor shaker (New Brunswick Scientific, Incubator Shaker, Model G25) for large scale cultures, or in a water bath shaker (New Brunswick Scientific, Gyrotory Water Bath Shaker, Model G76) for cultures of 50 ml or less. In order to obtain a clonal isolate of a bacterial population, overnight liquid cultures were streaked on LB agar plates, grown at 37°C overnight and were stored at 4°C. For long-term storage, 1 ml aliquots containing 10% glycerol (Fisher) were frozen at -20°C. Viable bacteria were recovered by thawing these cultures and by transferring a loop of bacteria to liquid medium.

3.4 Transformation

Transformation was carried out essentially as described by Sambrook *et al.*, (1989). One colony *E. Coli* of either DH5α or GM48, was grown overnight in sterile LB medium in a 37°C shaking water bath. Half of a ml of the overnight culture was inoculated into 50 ml of LB and grown to log phase (approximately 2.5 - 3 hours) and centrifuged at 3000 rpm (Centra-7R) for 15 minutes at 4°C. Cell pellets were resuspended in 25 ml of ice cold transformation buffer (75 mM $CaCl_2$; 5 mM Tris.HCl; pH 7.6) and incubated on ice at 4°C for four hours or overnight. Subsequently, the bacteria were collected by centrifugation at 3000 rpm for 15 minutes at 4°C and resuspended in 1 ml of ice cold transformation buffer. Approximately 10 ng of transforming DNA was added to 100 µl of bacteria and the cells were incubated on ice for 30 minutes with occasional shaking, then heat shocked for 2 minutes at 42°C and subsequently incubated on ice for 2 minutes. The cells were then diluted with 900 µl of warm SOC medium (2% bacto-tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$ and 20 mM glucose). Finally, serial dilutions were carried out and 100 µl of bacteria were plated on LB agar (2%) plates supplemented with 25-50 µg/ml ampicillin (Amp; Sigma) and 0.3 ug/ml Xgal (5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside, Sigma).

3.5 Plasmid DNA Preparations

3.5.1 Mini-Prep Plasmid DNA Preparations

Mini-prep plasmid DNA preparations were essentially carried out as described by Birnboim and Doly (1979). Colonies were picked from agar plates and grown in 3 ml of LB supplemented with Amp for 8-12 hours at 37°C in a water bath shaker . The bacterial cells were centrifuged for 30-45 sec at 14,000

rpm. The cell pellet was resuspended in 100 µl lysozyme solution (50 mM glucose; 10 mM EDTA; 24 mM Tris.HCl; pH 8) that contained 2 mg/ml lysozyme (Sigma) and incubated on ice for 30 minutes. Spheroplasts were then lysed and cellular proteins were denatured by adding 200 µl freshly prepared alkaline SDS (0.2 N NaOH; 1% SDS). After 10-15 min incubation at room temperature, 150 µl of sodium acetate (3M NaAc; pH 4.8) was added. Following a 30 min incubation on ice, the tubes were centrifuged at 14,000 rpm for 5 min. The clear supernatant fluid was then transferred to fresh tubes and ethanol precipitated using 2 volumes of -20°C 95% ethanol. The DNA pellet was resuspended in 300 µl distilled water and re-precipitated with 2 volumes of ethanol. The final pellet was dried, and resuspended in TE buffer (10 mM Tris, 1 mM EDTA; pH 7.5).

3.5.2 Large Scale Plasmid DNA Preparations

For large scale plasmid DNA purifications, cultures were grown to log phase. Extraction of plasmid DNA was essentially carried out according to the mini-prep procedure outlined earlier, except that 500 ml of bacteria culture was centrifuged at 4000 rpm for 50 min at 4°C and resuspended in 10 ml of lysozyme buffer, followed by the addition of 20 ml of alkaline SDS and 15 ml of 3M NaAC as described above. The final DNA pellet was dried and resuspended TE Buffer.

3.6 Restriction Endonuclease Analysis

The restriction enzymes used in this study [supplied by Promega, Boehringer Mannheim, or New England Biolabs (NEB)] were used according to suppliers recommendations. All restriction enzyme digestions were carried out at 37°C, for several hours or overnight. Typically, 10 µl plasmid DNA (20-30 ug) was mixed with 3 µl of 10X restriction buffer (provided by suppliers), 5 to 10 units of the restriction enzyme, 0.5 µl DNAse-free RNAse (Boehringer Mannheim) and

the volume was made up to 30 ul with distilled water. Restriction enzymes were inactivated after digestion by phenol: chloroform: isoamyl alcohol extraction, followed by ethanol precipitation. except when the reaction mixture was used for agarose gel electrophoresis in which case the samples were mixed with (10% v/v) stopper (20% glycerol; 2% SDS, 0.5% bromophenol blue), and immediately loaded into an agarose gel.

3.6.1 Gel Electrophoresis

DNA fragments of the recombinant plasmids were electrophoretically separated on 0.8 to 1.2 % agarose (Sigma) gels at 5 volts per cm in a Tris Acetate buffer (40 mM Tris, 57.1 ml/L glacial acetic acid, 1mM EDTA). The samples were loaded into submerged wells after the addition of 10% v/v stopper. During or after electrophoresis the gels were stained with ethidium bromide (0.5 μg/ml) and then placed on a UV transluminator and photographed with a mounted Polaroid land camera with Polaroid type 57 film. The sizes of the fragments were graphically estimated using λ- *Hin*dIII (Promega) as a molecular-size marker.

3.7 Radiochemicals

The radiochemicals used in this study were α[32P] dCTP; α[32P] dATP; and γ[32P] dATP (purchased from ICN), and α [35S] dATP (purchased from Dupont or ICN).

3.8 DNA:DNA Hybridization

3.8.1 Southern Blot

BAV2 DNA, extracted by the method of Hirt (1967) and plasmid DNA were digested with various restriction endonucleases. Ten to fifteen μg of each

DNA digest was then electrophoresed on 0.8 to 1.0 % agarose horizontal slab gels. Radioactively labeled λ- *Hin*dIII digest was used as a molecular -size marker. Gels were run at 30 to 50 volts overnight. The gel was immersed in 0.25N HCl and gently agitated for 8 to 10 min at room temperature. The gel was then rinsed briefly with deionized water. Next, the gel was immersed in 150 ml of a 1 M NaCl and 0.5 M NaOH solution and gently agitated for 15 min at room temperature twice. Soaking in a 0.5 M Tris pH 7.4 and 1.5 M NaCl solution twice for 15 minutes at room temperature neutralized the gel. The DNA was then transferred from the gels to nylon membranes (Gene Screen Plus, Dupont) according to the Southern transfer technique (Southern, 1975). The gel was then transferred. The transfer buffer was 6X SSC (diluted from 20X SSC 3M NaCl; 0.3M NaCitrate). After the transfer had occurred, the nylon filter was removed and was baked for 20 min to 2 h at 80 °C in a vacuum oven.

The recombinant plasmids were radioactively labeled using a random primer method developed by Feinberg and Vogelstein (1983; 1984). The Prime-a-Gene labeling kit supplied by Promega was used. Twenty five ng of denatured plasmid DNA was mixed with sterile water to a final volume of 25 µl. To the mixture 5 µl of 5X labeling buffer (250 mM Tris-HCl, pH 8.0; 25 mM MgCl2; 10 mM DTT; 1 M HEPES, pH 6.6; and 26 A260 units /ml random hexadeoxyribonucleotides), 1 µl of mixture of the non labeled dNTPs., 2 µl of nuclease-free BSA, 2.5 µl (25 µCi) [α32P] dCTP or [α32P] dATP and 5 units of polymerase (Klenow) were added. The mixture was incubated at room temperature for a minimum of 60 min, followed by a 2 min incubation at 95-100°C and subsequently chilled on ice. EDTA was added to a final concentration of 20 mM. If the labeled DNA was not used immediately, it was stored at -20°C and reheated at 95-100°C for 2 min and chilled on ice prior to hybridization.

A Klenow (Promega) filling reaction was used to label the λ-*Hin*dIII marker. Essentially, 1 µl of λ marker, 0.5 µl of each of the non labeled dNTPs, 4 µl of 10X reaction buffer (0.5 M Tris.Cl, pH 7.6, 0.1 M $MgCl_2$), 28.5 µl distilled water, 4 µl of (40 µCi) [α32P] dCTP or [α32P] dATP and 1µl of Klenow were added to make a final volume of 40 µl. After a 30 min incubation at 37°C, 20 µl of stopper was added prior to storage at -20°C.

The labeled recombinant plasmids were hybridized to the filters as described by Mahmoudi and Lin (1989). The plasmid pE3Dx1 used to localize the E3 region within the BAV2 genome was obtained from Dr. Y. Haj-Ahmad (Brock University, St. Catharines, Ontario, Canada). DNA-DNA hybridization was performed in a total volume of 20 ml at 68°C in the presence of 0.5 M $NaP0_4$, 1% BSA, 7% SDS, 0.04 M EDTA and 2 µg of carrier DNA. Prehybridization was carried out for 30 to 45 min at 68°C in a water bath shaker, followed by hybridization for 8 to 12 hours under the same conditions. The filter was washed twice for 15 min each with a 2X SSC, 0.1% SDS solution at 68°C, dried and exposed to Kodak X-Omat AR films.

3.8.2 Hybridization to DNA in Agarose Gels

Digested DNA fragments were run on 0.8 to 1.0 % agarose gels , stained with ethidium bromide and photographed. The gels were dried onto Whatman paper (Fisher) and stored dry at 4°C. Prior to hybridization, gels were rehydrated in distilled water for 5 min at room temperature (RT) , peeled away from the Whatman paper , denatured for 15 to 30 min at RT with 1.0 M NaCl, 0.5 M NaOH, and neutralized in 3 M NaCl, 0.5 M Tris-HCl pH 7.4 two times for 15 min each at RT. The Prehybridization was carried out at 37°C for 2 hours in 50 ml of 6X SSC and 0.5% NP40 plus 50 ml of solution A (25 ml formamide, 12.5 ml 20X SSC, 5 ml Denhardt's solution, 2.5 ml 1 M Tris pH 7.9, 1 ml of carrier DNA (10

μg/ml) and approximately 4 ml of distilled water). The gel was removed from the prehybridization solution and hybridization was carried out at 37°C overnight in 100 ml of hybridization solution (6X SSC and 0.5% NP40). The gel was then washed twice with 100 ml of 6X SSC for 10-20 min at RT, wrapped in plastic wrap and exposed on Kodak X-Omat film.

3.9 DNA Sequencing

Nucleotide sequencing was carried out according to the method of Sanger (1977) by direct plasmid sequencing . The Sequenase kit version 2.0 from U.S Biochemical (Cleveland , Ohio) was used for sequencing according to supplier's recommendations. Usually, the inserts were sequenced from both ends using the M13 universal primer and/or M13 reverse primer. However, when plasmids were too large to be sequenced through and convenient restriction enzyme sites were not available for subcloning, synthesized oligonucleotides were used (Vetrogen Corp., London, On.).

Plasmid DNA was prepared by the alkaline lysis procedure described earlier followed by RNase treatment at 37°C for 30 min The DNA was denatured by adding 0.1 volumes of 2 M NaOH and 2 mM EDTA and incubating at 37°C for 30 min (Lim & Pene, 1988). The mixture was neutralized by adding 0.1 volumes of 3M sodium acetate (pH 4.6) and then the DNA was precipitated with 2-4 volumes ethanol, dried and resuspended in distilled water. The DNA concentration was determined using a spectrophotometer.

Approximately 0.5 picomoles of a 16-mer oligonucleotide primer was annealed to 3-5 μg of denatured DNA in the presence of 1X Sequenase buffer (200 mM Tris-HCl (pH 7.5), 100 mM $MgCl_2$ and 250 mM NaCl). The volume was adjusted to 10 μl with distilled water. The mixture was heated to 65 °C for 2 min

then allowed to cool to room temperature over a period of about 30 min. The mixture was stored on ice.

For the labeling reaction, 6 μM of DTT (dithiothreitol), 0.2 μM each dNTP (d GTP, dTTP, dCTP and [α-^{35}S or α-^{32}P] dATP)and 3.25 units of polymerase enzyme (Sequenase version 2.0) were added to the ice- cold DNA mixture. Following extension at room temperature for 2-5 min, 3.5 μl samples were added to each of the four tubes containing 2.5 μl of one of the dideoxynucleotide (8 μM ddNTP), 80 μM of each dNTP and 50 mM NaCl. After 5 min at 37 °C, the reactions were terminated using 4 μl of stop solution (95 % formamide, 20 mM EDTA, 0.5 % bromophenol blue and 0.05% xylene cyanol FF) and stored at -20°C.

The sequencing gel solution [6 % Long Ranger (J.T. Baker)], 8 M urea, and 1 X TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA ; pH 8.0) was filtered using # 1 Whatman paper. Two hundred and fifty μl of 10 % ammonium persulfate and 25 μl of TEMED (N,N,N',N'- tetramethylethlene diamine; Bio-Rad, ICN) were added and the gel was poured and allowed to polymerize for 1 hour. The sequencing reaction mixtures were heated for 2 min at 75 -80 °C immediately prior to loading. The samples (2.5 -3 μl) were loaded into the sharktooth well on the upper gel surface. Electrophoresis was carried out at a constant current of 1.25 amp/cm in 0.6 X TBE buffer for 2 to 7 hours. The gel was removed from the glass plate by adsorption to Whatman paper (#3), covered with plastic film, and vacuum dried (Bio-Rad Model 583 Gel Dryer). The dried gel was exposed to Kodak X-OMAT AR X-ray film for 24 hours (for ^{32}P labeled samples) or 72 h (for ^{35} S labeled samples).

DNA sequences were recorded and analyzed using the Macintosh program DNA Strider version 1.0. Homology searches between the predicted BAV2 and known adenovirus DNA and protein sequences were done using the

European Molecular Biology Laboratory (EMBL) nucleic acid and protein sequence data base. The FASTA program was employed (Pearson & Lipman,1988).

CHAPTER 4: RESULTS

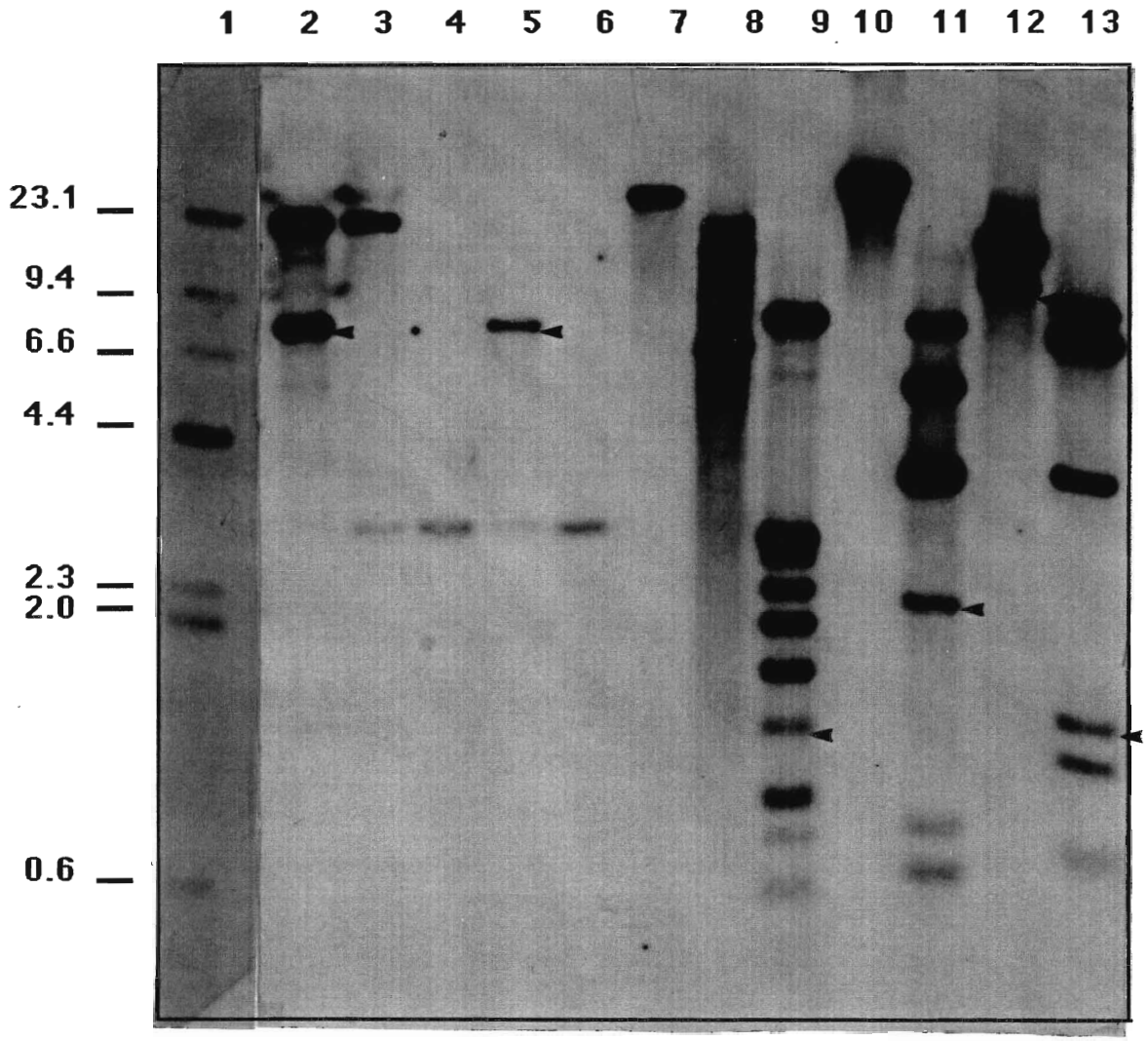
This study was multifaceted; each phase depended on the success of the previous step. The starting material was the plasmid pEA48 which consisted of the BAV2 *Eco*RI A fragment, believed to contain the E3 region. First, hybridization techniques were used to determine that the DNA insert was viral in origin. Secondly, a detailed restriction enzyme map of pEA48 needed to be constructed. Thirdly, hybridization techniques were used to localize the E3 region within the insert using an Ad5 plasmid consisting of part of the two regions encoding the protein pVIII and the fibre that flank the E3. Finally, the E3 and neighbouring regions were sequenced. The resulting sequence was compared to known adenoviral DNA and protein sequences.

4.1 DNA:DNA Hybridization Analysis

The *Eco*RI plasmids pEA48, pEB*, pEB113 and pED generated by Kirsty Salmon (Brock University, St. Catharines, Ontario, Canada) were constructed from Hirt (1967) extracted DNA that was cut with *Eco*RI and ligated to pUC19. Consequently, cell genomic DNA was also present in the ligation reaction and it was unclear whether the *Eco*RI inserts were of viral origin. A ^{32}P-labeled pEA48 probe was hybridized to the *Eco*RI recombinant plasmids (pEA48, pEB*, pEB113 and pED constructed by Kirsty Salmon) digested with *Eco*RI and BAV2 DNA digested with *Eco*RI,*Bam*HI, *Cla*I, *Hin*dIII, *Kpn*I, *Pst*I, *Sal*I and *Xba*I. The DNA fragments that hybridized to the ^{32}P- labeled pEA48 probe appear as

Figure 4 :

An autoradiogram of a hybridization blot of the*Eco*RI plasmids (pEA48, pEB*, pEB and pED) digested with *Eco*RI and BAV2 DNA digested with *Eco*RI,*Bam*HI, *Cla*I, *Hin*dIII, *Kpn*I, *Pst*I, *Sal*I, and *Xba*I, respectively (lanes 2, 7-13, respectively) hybridized to a pEA48 probe. Lane 1 represents λ -*Hin*dIII marker in kbp, and lanes 3-6 represent pEA48, pEB*, pEB, and pED, respectively This hybridization blot was done on a washed nylon filter that was previously probed with a pEB probe. The bands that are not the result of the hybridization with the pEA48 probe were determined from other autoradiograms and are indicated with a arrow. The proposed plasmids pEB* and pED were determined to be genomic in origin.


1 2 3 4 5 6 7 8 9 10 11 12 13
23.1
9.4
6.6
4.4
2.3
2.0
0.6

bands (fig. 4). The sizes of the DNA fragments correspond to the sizes of pEA48 cut with the same enzymes (fig. 7). It was determined from other autoradiograms that some of the bands were residual from a previous hybridization with ^{32}-P labelled pEB113 and those bands were indicated with an arrow. The plasmids pEB* and pED did not hybridize to BAV2 and the DNA inserts were determined to be genomic in origin (data not shown).

A ^{32}P-labeled pEB113 probe was hybridized to the *Eco*RI plasmids (pEA48, pEB*, pEB113 & pED) that were digested with *Eco*RI and to BAV2 DNA digested with *Eco*RI,*Bam*HI, *Cla*I, *Hin*dIII, *Kpn*I, *Pst*I, *Sal*I and *Xba*I. The DNA fragments that hybridized to the ^{32}P- labeled pEB113 probe appear as bands (fig. 5). The sizes of the DNA fragments were determined using the λ *Hin*dIII digested marker (Fig. 5, lane I). The DNA insert of pEB113 co-migrated with the B fragment of BAV2 digested with *Eco*RI (fig. 5, lane 5 & lane 2, respectively). The approximate size of this fragment determined from agarose gels (dat not shown) is 6.8 kbp. The size of the resulting *Bam*HI DNA fragment was too large to estimate (fig. 5, lane 7). When *Cla*I digested BAV2 was probed with ^{32}P-labeled pEB113 only one band was generated (fig, 5, lane 8). *Hin*dIII digested BAV2 generated 4 DNA fragments (7.0 kbp, 2.6 kbp, 1.3 kbp, .5 kbp; fig.5, lane 9). The band that is 5.9 kbp is the result of a partially digested BAV2 (fig. 5, lane 9). *Kpn*I did not digest BAV2; the resulting band is uncut viral DNA (fig. 5, lane 10). *Pst*I digestion resulted in 3 DNA fragments (3.2 kbp, 2 kbp & 0.6 kbp; fig. 5, lane 11). The largest band (5.5 kbp) is the result of partially digested BAV2. The sizes of the *Sal*I fragments are 12.1 kbp and 8.5 kbp. Finally, 3 DNA fragments (7.6 kbp, 6.4 kbp, & 1.3 kbp) were generated when BAV2 was digested with *Xba*I (fig. 5 lanes 12, 13, repectively).

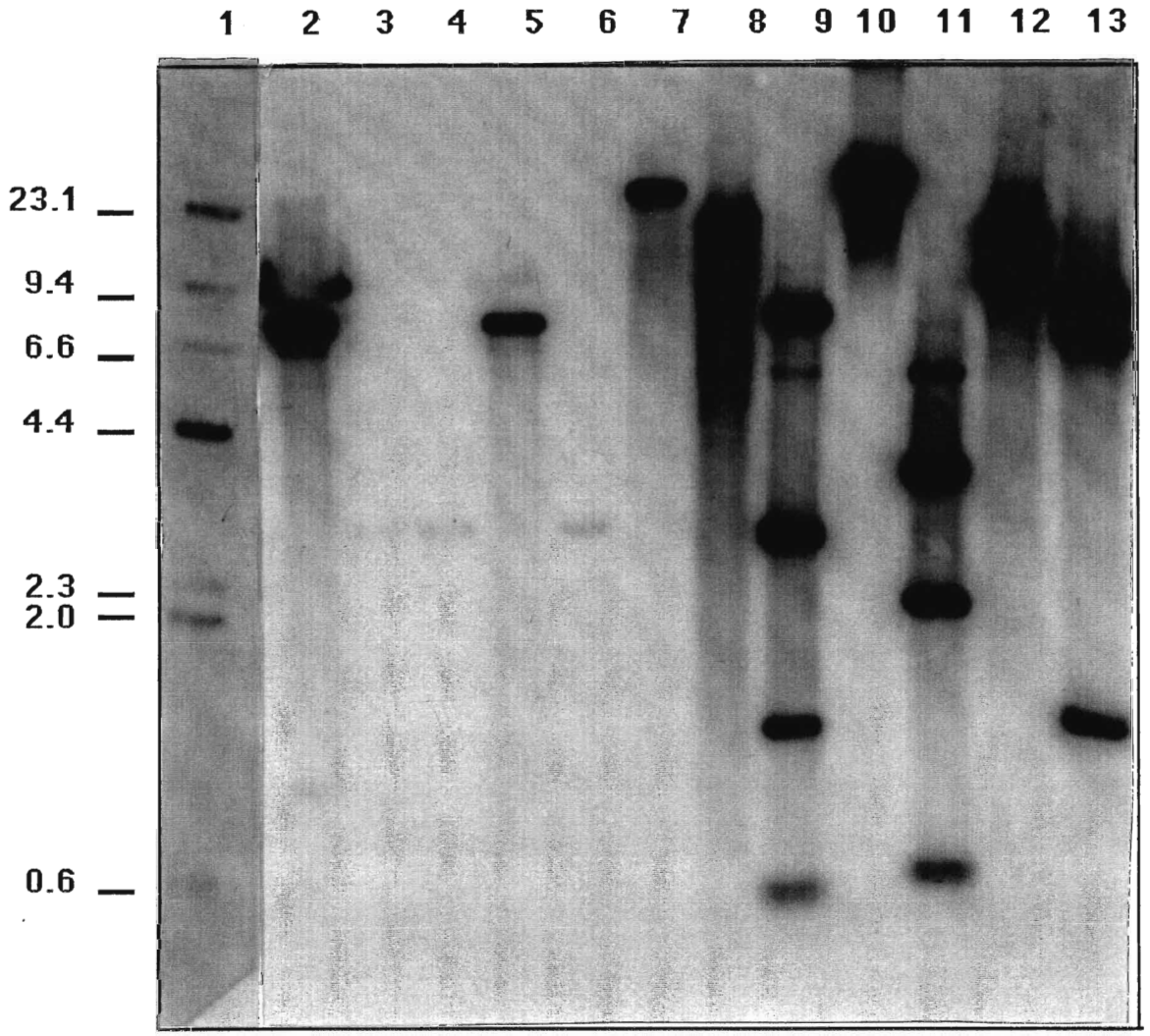


Figure 5 :

An autoradiogram of a hybridization blot of the *Eco*RI plasmids (pEA, pEB*, pEB113 and pEC) digested with *Eco*RI and BAV2 DNA digested with *Bam*HI, *Cla*I, *Hin*dIII, *Kpn*I, *Pst*I, *Sal*I, and *Xba*I, respectively (lanes 7-13) hybridized to a pEB113 probe. Lane 1 represents λ -*Hin*dIII marker (the size is indicated in kbp on the side) , Lane 2 represents BAV2 digested with *Eco*RI and lanes 3-6 represent pEA48 ,pEB*, pEB113, and pED.

4.1.1 Restriction Enzyme Analysis of pEA48

The plasmid pEA48 (grown in GM48) containing the *Eco*RI A fragment of BAV2 was digested with the following enzymes: *Bam*HI, *Cla*I, *Eco*RI, *Hin*dIII, *Kpn*I, *Not*I, *Pst*I, *Sal*I, *Xba*I and *Xho*I (fig. 6A). The resulting fragments were electrophoretically separated on a 1.2 percent agarose gel along with a sample of a molecular size marker (λ phage DNA digested with *Hin*dIII). The sizes of the fragments were interpolated from the standard curve constructed from the distance migrated versus known size standard. This technique is believed to be accurate within 100 bp (Sambrook *et al.,* 1989). The size estimates of DNA fragments that are larger than 9.4 kbp are questionable and if possible were determined by the summation of smaller fragments. For example, the size of pEA48 was not determined by the migration of the *Eco*RI DNA fragment alone but rather by adding up the smaller DNA fragments of pEA48 digested with *Hin*dIII and/or *Pst*I. All of the DNA fragments generated by both of these enzymes were under 9.4 kbp. The sizes of the DNA fragments were not determined by a single gel but rather each value represents an average molecular weight from 4 to 5 gels.

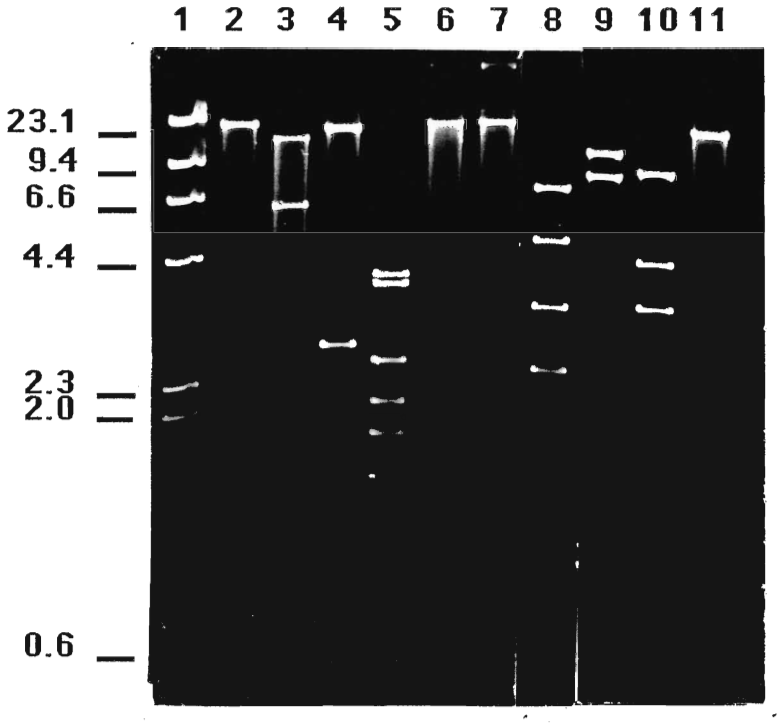
Two of the ten restriction enzymes (*Bam*HI & *Kpn*I) that were used did not cut the DNA insert. *Bam*HI and *Kpn*I generated a single band from a cut within the multiple cloning site (MCS) of pUC19. The remaining 8 restriction enzymes that were used appear to cut the DNA insert, and the approximate location of the restriction sites are indicated in the physical map (fig. 6B). *Cla*I produced 2

Figure 6:

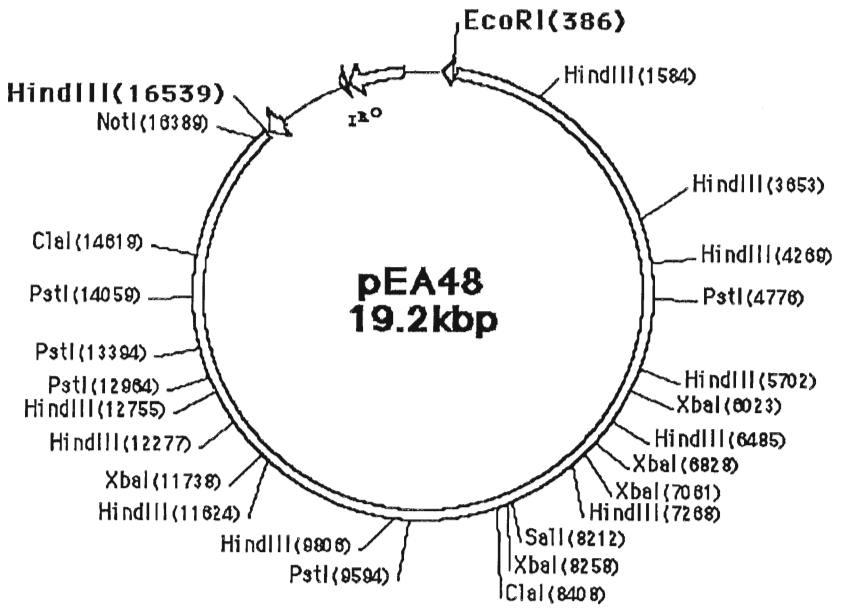
A: Photograph of an agarose gel of pEA48 digested with 10 restriction enzymes. Lane 1, λ marker; lanes 2-11pEA48 digested with *Bam*HI, *Cla*I, *Eco*RI, *Hind*III, *Kpn*I, *Not*I, *Pst*I, *Sal*I, *Xba*1 and *Xho*I, respectively.

B: Physical map of pEA48 . Restriction enzyme recognition sites have been mapped around the plasmid.

A


1 2 3 4 5 6 7 8 9 10 11
23.1
9.4
6.6
4.4
2.3
2.0
0.6

B


EcoRI(386)
HindIII(1584)
HindIII(3653)
HindIII(4269)
PstI(4776)
HindIII(5702)
XbaI(6023)
HindIII(6485)
XbaI(6828)
XbaI(7061)
HindIII(7268)
SalI(8212)
XbaI(8258)
ClaI(8408)
PstI(9594)
HindIII(9806)
HindIII(11624)
XbaI(11738)
HindIII(12277)
HindIII(12755)
PstI(12964)
PstI(13394)
PstI(14059)
ClaI(14619)
HindIII(16539)
NotI(16389)
IR
pEA48
19.2kbp

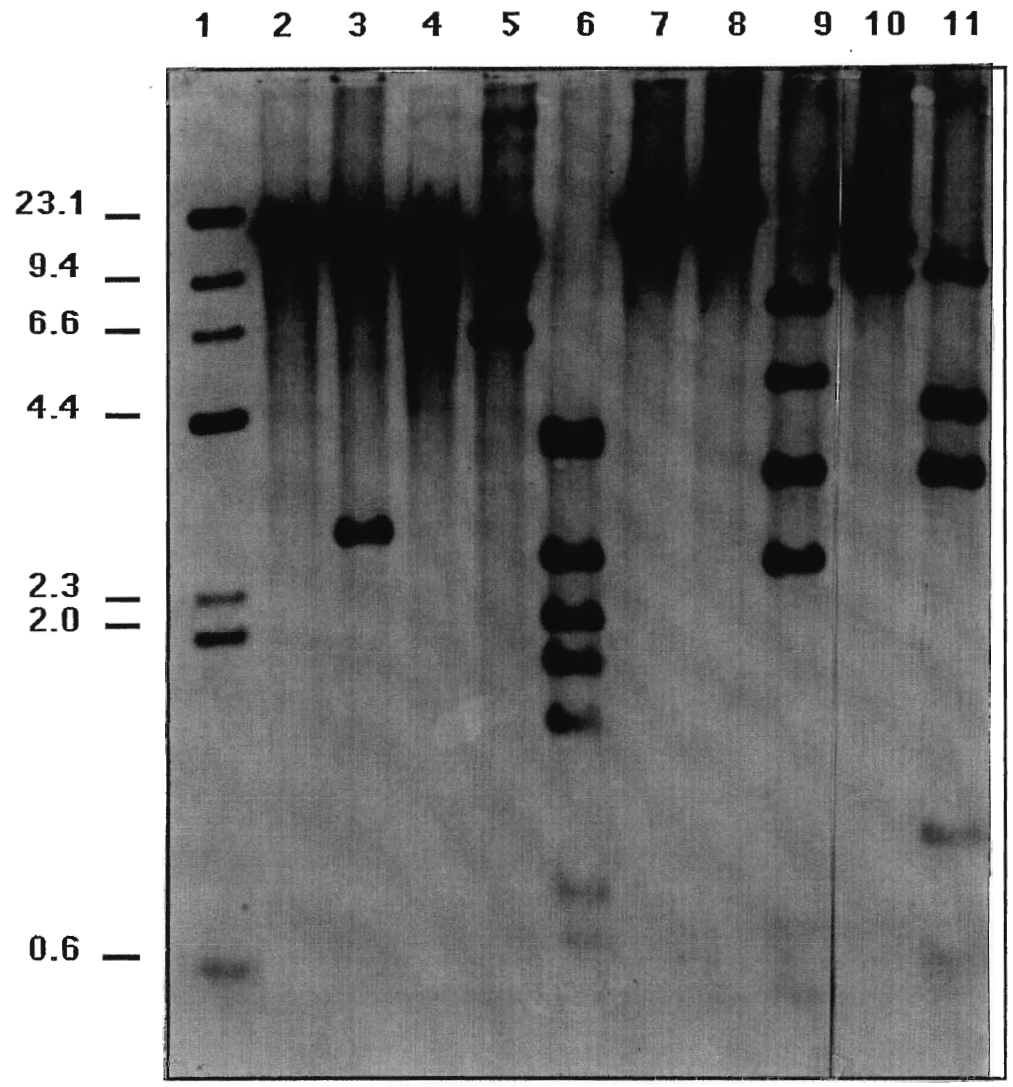


Figure 7:

An autoradiogram of a hybridization blot of BAV2 digested with *Eco*RI and pEA48 digested with *Eco*RI, *Bam*HI, *Cla*I, *Hin*dIII, *Kpn*I, *Not*I, *Pst*I, *Sal*I, and *Xba*I, respectively (lanes 3-13) hybridized to a pEA48 probe. Lane 1 represents λ -*Hin*dIII marker (in kbp on the side) and lane 2 represents BAV2 digested with *Eco*RI.

visible bands with sizes of 12.3 kbp and 6.0 kbp on the agarose gel (fig. 6,lane 3) and the autoradiogram (fig. 7, lane 4). The autoradiogram was generated from a Southern blot of ^{32}P -labelled pEA48 hybridized to pEA48 digested with the following enzymes: *Bam*HI, *Cla*I, *Eco*RI, *Hin*dIII, *Kpn*I, *Not*I, *Pst*I, *Sal*I, *Xba*I and *Xho*I. This allowed for the detection of small DNA fragments.

*Eco*RI generated two bands (fig. 6A , lane 4; fig. 7, lane 3); a large fragment approximately 16 kbp and pUC19, 2.7 kbp. The large DNA fragment comigrated with the large BAV2 *Eco*RI fragment (fig. 7, lane 2, 3). When pEA48 was digested with *Hin*dIII, 10 DNA fragments were generated (3.9* kbp, 3.7 kbp, 2.5 kbp, 2.1 kbp, 1.8 kbp, 1.5 kbp, 0.8 kbp, 0.8 kbp,0.7 kbp, 0.5 kbp). Although the smaller fragments (0.8 kbp, 0.8 kbp,0.7 kbp, 0.5 kbp) were not apparent in figure 6A they were detected on other agarose gels. The DNA fragments that contain pUC19 DNA are indicated by an asterisk. The fragments containing pUC19 were determined by double digests with a given enzyme and *Eco*RI. The actual size of the Ad portion of the DNA fragment was determined by subtracting 2.7 kbp (the size of pUC19) from the size apparent on the gel. *Not*I and *Xho*I digestions generated only 1 band each. Unlike *Bam*HI and *Kpn*I, these cuts were within the DNA insert and not in the plasmid pUC19. *Not*I and *Xho*I do not cut the plasmid vector, pUC19. Digestion with *Pst*I generated six fragments (6.9* kbp, 4.7 kbp, 3.3 kbp, 2.4 kbp, 0.7 kbp, and 0.4 kbp). The two smallest fragments can be seen clearly on the autoradiogram (fig. 5, lane 9). Of the two remaining restriction enzymes, *Sal*I (fig. 4, lane 10 & fig. 5, lane 11) generated two fragments (10* kbp, & 7.0 kbp) and *Xba*I generated 6 DNA fragments (8.3* kbp, 4.3 kbp, 3.5 kbp, 1.1 kbp, 0.8 kbp & 0.3 kbp). Figure 6B is the physical map of pEA48. Locations of restriction enzyme sites were determined by partial and multiple digestions.

4.2 Localization of the E3 Region

Southern hybridization techniques were used to locate the E3 region in the BAV2 genome. The plasmid pE3dx1 (obtained from Dr. Haj-Ahmad, Brock University, St. Catharines, Ontario, Canada) containing part of the E3 region of Ad5 was used as a probe to identify the BAV2 DNA fragments that contain the E3 region. Figure 8 is a schematic representation of the Ad5 E3 probe used.

The resulting DNA fragments of BAV2 DNA digested with *Eco*RI, *Cla*I, *Hin*dIII, *Kpn*I, *Pst*I, *Sal*I, and *Xba*I, that hybridized to ^{32}P-labeled Ad5 pE3dx1 probe, are shown in figure 9. The sizes of the DNA fragments were determined using λ-*Hin*dIII digested marker (fig. 9, lane 1). The location in the BAV2 genome and the estimated size of the DNA fragment that hybridized to the Ad5 pE3dx1 probe are summarized in figure 10 . Only one band was visualized in each lane. The probe did hybridize to the *Eco*RI A fragment. The band seen in lane 5 is uncut BAV2 DNA. *Kpn*I does not digest the viral genome. The sizes of the BAV2 DNA fragments cut with *Cla*I, *Hin*dIII, *Pst*I, *Sal*I and *Xba*I that hybridized to the Ad5 pE3dx1 probe were approximately 12.0, 1.4, 4.8, 11.6, and 6.5 kbp respectively. The *Hin*dIII digest of BAV2 DNA generated several small fragments. The probe hybridized to a single fragment located between 74.8 and 79.2 mu. The location of this fragment within the BAV2 genome and the fact that the pVIII is highly conserved suggests that the hybridization to the probe was due to homology between the Ad5 pVIII and not the fibre protein. Assuming that the genetic organization of Ads is conserved, the probable start of the E3 region is likely to lie somewhere within this 1.4 Kbp fragment. As a consequence, DNA sequencing started at 74.8 mu.

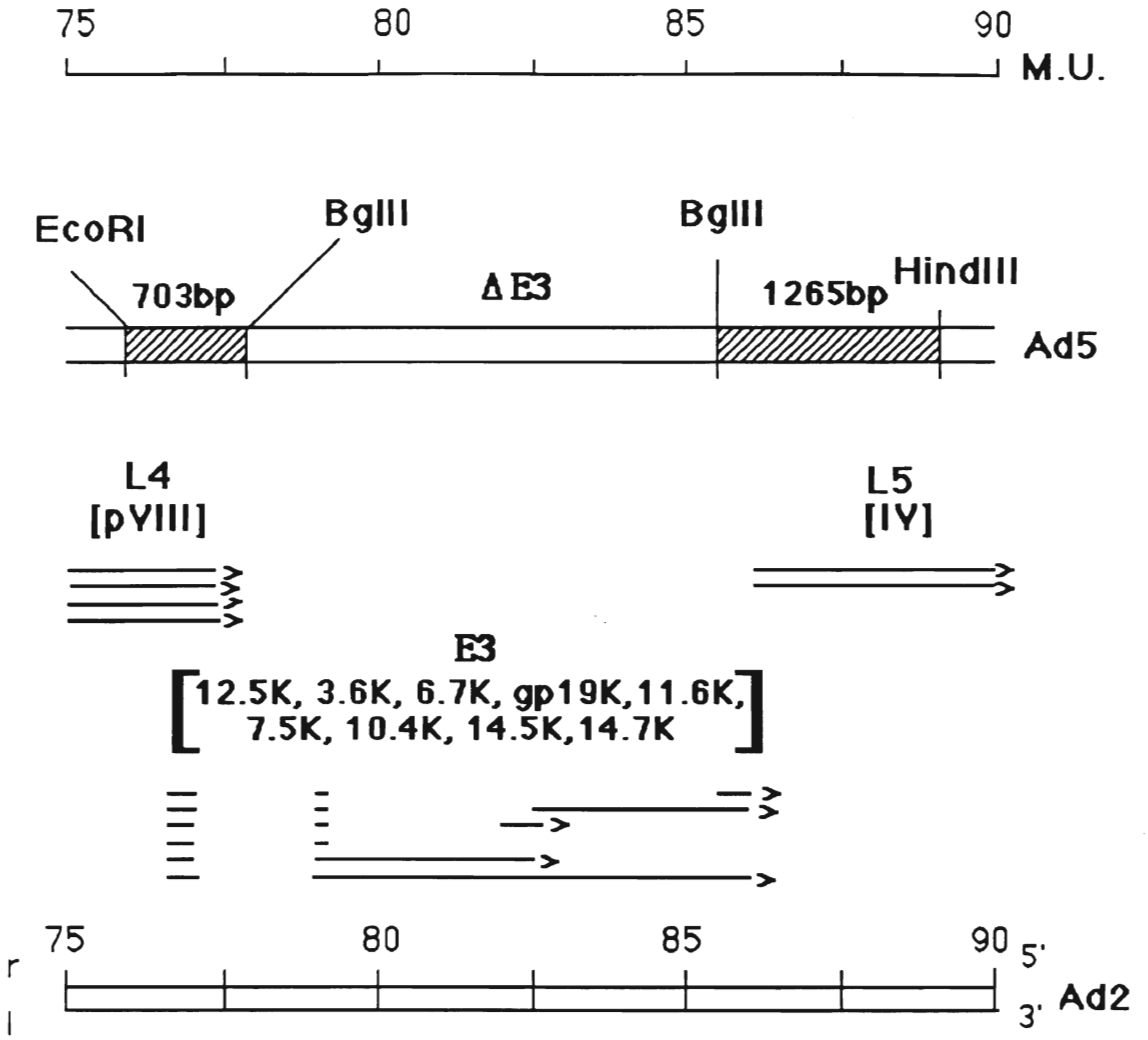


Figure 8:

A schematic drawing of the Ad5 sequences contained within pE3dx1 (Haj-Ahmad, 1986) probe used in the Southern hybridization studies to localize the E3 region in BAV2. The DNA between the two *Bgl*II sites was excised and the remaining DNA fragments are indicated by the cross bars. A modified schematic diagram of the Ad2 transcription unit map (reproduced from Wold & Gooding, 1991) was used to illustrate the location of the DNA sequences of the Ad5 pE3dxI plasmid within the Ad2 genome.

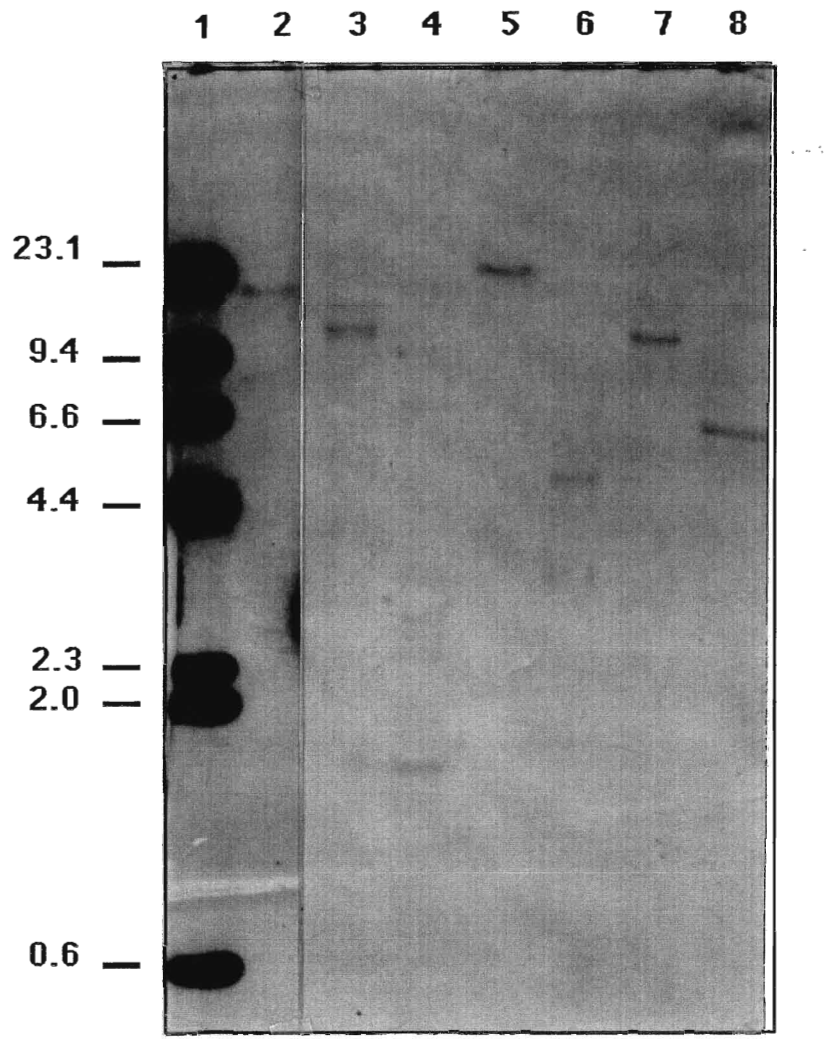


Figure 9:

Localization of the E3 region. BAV2 DNA digested with*Eco*RI, *Cla*I, *Hin*dIII, *Kpn*I, *Pst*I, *Sal*I, and *Xba*I, respectively (lanes 2-8), hybridized to the Ad5 pE3dx1 probe (Haj-Ahmad, 1986). Lane 1 represents λ - *Hin*dIII marker (in kbp on the side).

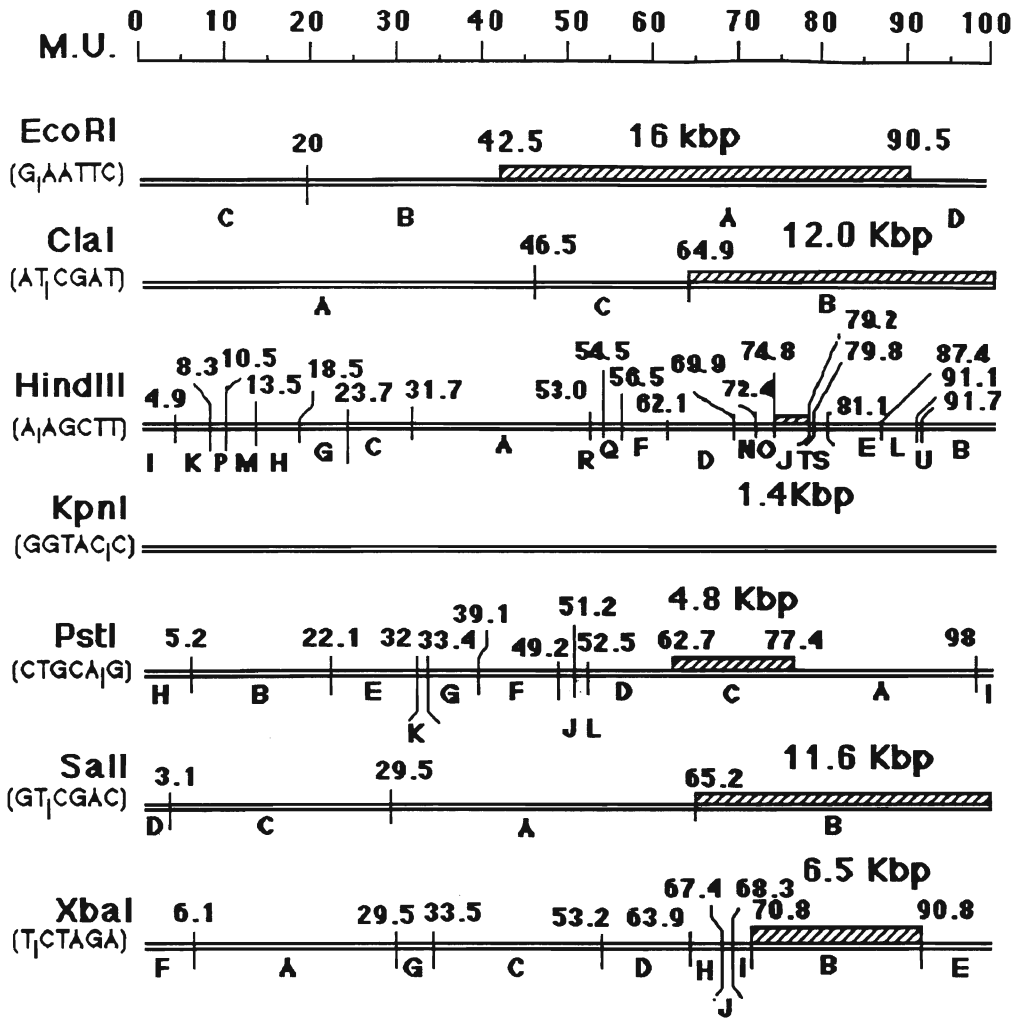


Figure 10:

A schematic representation of 7 restriction endonuclease nuclease maps of BAV2 DNA. Each map is divided into 100 map units, where 1 map unit is equal to 325 bp (constructed in collaboration with Kirsty Salmon, Brock Univesity, St. Catharines, Ontario, Canada). The DNA fragments that hybridized to the ^{32}P labelled pE3Dx1 probe are indicated with cross bars. The sizes of the fragments are indicated in kbp.

4.3 Sequencing the E3 and Neigbouring Regions

4.3.1 Sequencing Strategy

In order to facilitate the sequencing of the E3 region it was necessary to subclone fragments from pEA48. Cloning was carried out using a shot gun strategy outlined in figure 11. The plasmid name indicates the enzyme used (where p stands for plasmid, E, H, P, X for *Eco*RI, *Hin*dIII, *Pst*I, and *Xba*I respectively, and a number for lab notation). Also, custom synthesized oligonucleotides were used to sequence overlaps and regions that were difficult to subclone.

Figure 12 is the resulting nucleotide sequence from 74.8 to 90.5 map units of BAV2 corresponding to the E3 region and flanking sequences, together with the deduced amino acid sequences for the largest open reading frames (ORF). Within the 5375 bp fragment there were 17 ORFs in the r strand capable of coding for polypeptides of 50 or more residues (fig. 13). Four short ORFs for unidentified proteins were encoded in the complementary strand (l-strand). They were located between nucleotides 101 - 322, 101 - 406, 5029 - 5311 and 5029 - 5265.

ORFs 1, 4, 5, 6, 10, 14 have theoretical coding capacities of 16.8, 7.1, 14.6, 11.7, 37.1 and 59.6 KDa, respectively. The TESTCODE algorithm (Fickett, 1982) was used to predict whether an ORF is protein coding or noncoding. The prediction of whether these ORFs coded for proteins is summarized in table 1. ORFs 4 and 6, which encoded for polypeptides that were 7.1 and 11.7 KDa were predicted to code for proteins. The probability for E3 37.1 KDA encoded by ORF 10 is 0.77; too low to make a prediction. ORFs 5 was predicted to be noncoding. A homology search of the nucleotide and predicted amino acid sequences using the European Molecular Biology Laboratory (EMBL) nucleic acid and protein data base and the FASTA program (Pearson & Lipman, 1988), revealed that

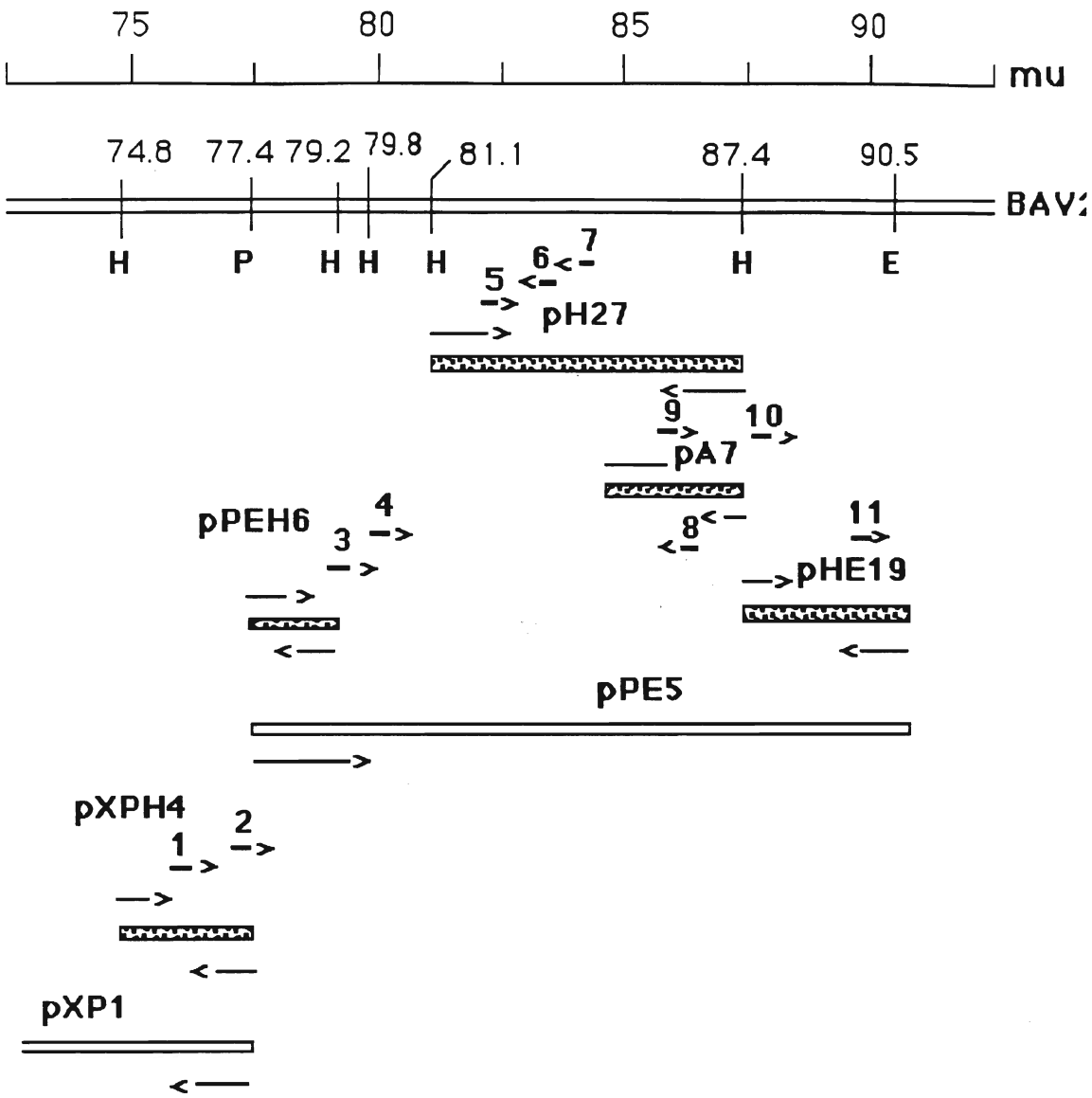


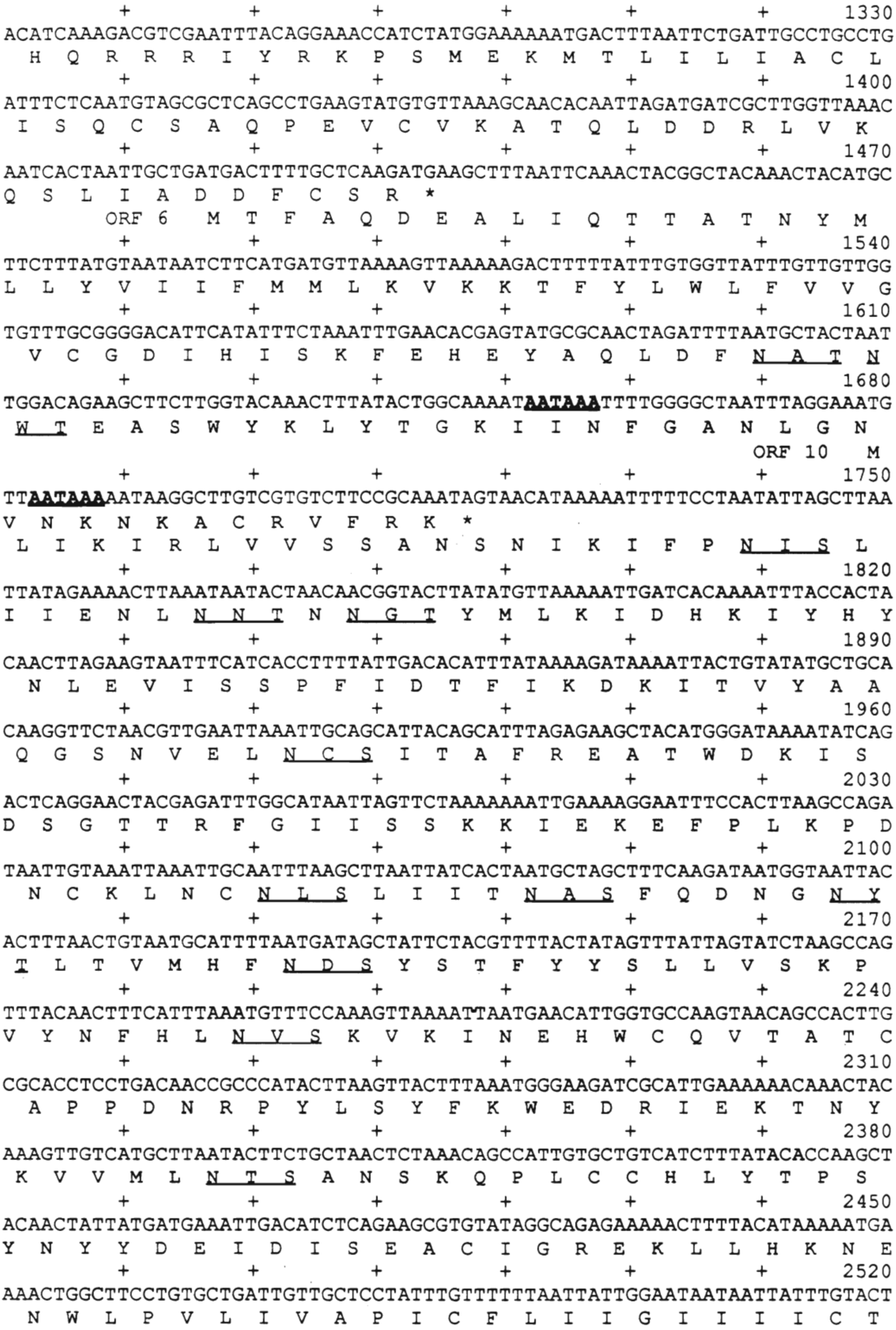
Figure 11 :

A schematic representation of the sequencing strategy used. The graph is in map units; each vertical bar represents 2.5 map units and each map unit is 325 bp. The horizontal bars indicate the recombinant plasmids used and their relative position. The arrows indicate the direction of sequencing. The arrows with numbers indicate the location and the direction of the synthesized oligonucleotides used. The letters represent the restriction enzyme sites; E- *Eco*RI; H- *Hin*dIII; P- *Pst*I

Figure 12:

The nucleotide sequence of BAV2 DNA, from 74.8 to 90.5 map units. The deduced amino acid sequence for the pVIII protein, the putative E3 proteins and the fibre protein, are shown in the single letter amino acid code. Potential N glycosylation sites (N-X-T/S) and polyadenylation signals (AATAAA) are underlined. The nucleotide sequence was analyzed using the program DNA Strider. ORFs and their locations are also indicated.

```
        +         +         +         +         +         +        70
AAGCTTGCCTACTTATAGTTCAGCATGAGCAAAGAAATACCCACGCCCTACATGTGGACTTTCCAGCCTC

        +         +         +         +         +         +       140
AATTAGGGCAAGCCGCAGGGCTTCCCAAGATTACTCAACCAGAATGAATTGGTTGAGCGCTGGACCGTCA

        +         +         +         +         +         +       210
ATGATTAGCCAGGTGGACTCTATCCGGGCGGAACGATAGAATTTTATTGCACCAGGCGGCAGTTTCGGCG

        +         +         +         +         +         +       280
ACGCTAGAGCAATCATGAACCCGCCTGCTTGGCCAGCCAGCCGCCTGTTCCAACCTGTGGATACGGTGCA
       ORF 1  M   N   P   P   A   W   P   A   S   R   L   F   Q   P   V   D   T   V   Q
        +         +         +         +         +         +       350
AACTTTGGAATTGCCACGTAACGAATTACTCGAAACTGCCATGACTAACTCTGGAATGCAGTTAGCCGGG
  T   L   E   L   P   R   N   E   L   L   E   T   A   M   T   N   S   G   M   Q   L   A   G
        +         +         +         +         +         +       420
GGAGGAGTGCACAGAACTAAAGATATAAAACCAGAAGACTTGGTCGGACGCGGCATTCAGCTCAACAGCT
 G   G   V   H   R   T   K   D   I   K   P   E   D   L   V   G   R   G   I   Q   L   N   S
        +         +         +         +         +         +       490
ACCAGCCGCCCACAACGCGGTTGAGGCCAGACGGAGTGTTTCAGCTAGCTGGAGGAAGCCGCTCTTCTTT
Y   Q   P   P   T   T   R   L   R   P   D   G   V   F   Q   L   A   G   G   S   R   S   S   F
        +         +         +         +         +         +       560
TAACCCCAGCATTAACACTCTGCTAACTCTTCAACCTGCTGCTTCTGTGCCGCGCTCCGGTGGCATCGGA
  N   P   S   I   N   T   L   L   T   L   Q   P   A   A   S   V   P   R   S   G   G   I   G
        +         +         +         +         +         +       630
GAAGTGCAATTTGTGCACGAGTTTGTGCCTTCTGTTTACTTCCAACCTTTCTCAGGACCGCCTGGATCAT
 E   V   Q   F   V   H   E   F   V   P   S   V   Y   F   Q   P   F   S   G   P   P   G   S
        +         +         +         +         +         +       700
ACCCTGACGAGTTCATCTACAACTTTGACGTTGCGACGGACTCTATTGACGGGTATGCATAACTGCAACT
Y   P   D   E   F   I   Y   N   F   D   V   A   T   D   S   I   D   G   Y   A   *
                                                  ORF 4   M   H   N   C   N
        +         +         +         +         +         +       770
CTGAGGACACCAGATTGTCTATAAAGGCGCTGTGCAAACTCCACTACCAATACTGCGCGTCAGACAACTG
S   E   D   T   R   L   S   I   K   A   L   C   K   L   H   Y   Q   Y   C   A   S   D   N   C
        +         +         +         +         +         +       840
CTTCTGGAAATCCAACTTGCTGCCATCATACTTCTGTGTTCTACCTGAAGCGCTTGAGGGAGACTGCGTG
  F   W   K   S   N   L   L   P   S   Y   F   C   V   L   P   E   A   L   E   G   D   C   V
        +         +         +         +         +         +       910
CCTGACTCGCTTCAGCCTGGACAGCCTACGCCTTGAAATCCCTCACTGCTATTTCAGCTACATTGCTAAC
 P   D   S   L   Q   P   G   Q   P   T   P   *
        +         +         +         .   +         +         +       980
TTTGATGGCCATGAGATCATCTGCAGCAGCAACCTCAAAAGCGGCTATGTTAAAGCCACTTGTCGGTGTA

        +         +         +         +         +         +      1050
GTGTCCAGGCGCACACTTGGCCCTTTTTCAAGGTATATGCTGGCTATTTAACAAACGTATGGTTCCTGTG
                                                                ORF 5
        +         +         +         +         +         +      1120
GATGATTCAAACACCGACTCTTGTTAGTGCTTCTAACTCTACAACCACCAACAAAACTTCATGCTTTAGC
  M   I   Q   T   P   T   L   V   S   A   S   N   S   T   T   T   N   K   T   S   C   F   S
        +         +         +         +         +         +      1190
AGCAATGACCCTACCTGCCGCTTAATCCAAGTATTCACAATTGGATTGCTTGGACTGCTGTGGTTTATAG
 S   N   D   P   T   C   R   L   I   Q   V   F   T   I   G   L   L   G   L   L   W   F   I
        +         +         +         +         +         +      1260
CTTTTGGGCTATGCTTTTGGCGCTTCCTACAATTTCGAGCCGAACAACAACTCCAAAAACAAGAAGATAA
A   F   G   L   C   F   W   R   F   L   Q   F   R   A   E   Q   Q   L   Q   K   Q   E   D   K
```


1330
ACATCAAAGACGTCGAATTTACAGGAAACCATCTATGGAAAAAATGACTTTAATTCTGATTGCCTGCCTG
H Q R R R I Y R K P S M E K M T L I L I A C L
1400
ATTTCTCAATGTAGCGCTCAGCCTGAAGTATGTGTTAAAGCAACACAATTAGATGATCGCTTGGTTAAAC
I S Q C S A Q P E V C V K A T Q L D D R L V K
1470
AATCACTAATTGCTGATGACTTTTGCTCAAGATGAAGCTTTAATTCAAACTACGGCTACAAACTACATGC
Q S L I A D D F C S R *
ORF 6 M T F A Q D E A L I Q T T A T N Y M
1540
TTCTTTATGTAATAATCTTCATGATGTTAAAAGTTAAAAAGACTTTTTATTTGTGGTTATTTGTTGTTGG
L L Y V I I F M M L K V K K T F Y L W L F V V G
1610
TGTTTGCGGGGACATTCATATTTCTAAATTTGAACACGAGTATGCGCAACTAGATTTTAATGCTACTAAT
V C G D I H I S K F E H E Y A Q L D F N A T N
1680
TGGACAGAAGCTTCTTGGTACAAACTTTATACTGGCAAAATAATAAATTTTTGGGGCTAATTTAGGAAATG
W T E A S W Y K L Y T G K I I N F G A N L G N
ORF 10 M
1750
TTAATAAAAATAAGGCTTGTCGTGTCTTCCGCAAATAGTAACATAAAAATTTTTCCTAATATTAGCTTAA
V N K N K A C R V F R K *
L I K I R L V V S S A N S N I K I F P N I S L
1820
TTATAGAAAACTTAAATAATACTAACAACGGTACTTATATGTTAAAAATTGATCACAAAATTTACCACTA
I I E N L N N T N N G T Y M L K I D H K I Y H Y
1890
CAACTTAGAAGTAATTTCATCACCTTTTATTGACACATTTATAAAAGATAAAATTACTGTATATGCTGCA
N L E V I S S P F I D T F I K D K I T V Y A A
1960
CAAGGTTCTAACGTTGAATTAAATTGCAGCATTACAGCATTTAGAGAAGCTACATGGGATAAAATATCAG
Q G S N V E L N C S I T A F R E A T W D K I S
2030
ACTCAGGAACTACGAGATTTGGCATAATTAGTTCTAAAAAAATTGAAAAGGAATTTCCACTTAAGCCAGA
D S G T T R F G I I S S K K I E K E F P L K P D
2100
TAATTGTAAATTAAATTGCAATTTAAGCTTAATTATCACTAATGCTAGCTTTCAAGATAATGGTAATTAC
N C K L N C N L S L I I T N A S F Q D N G N Y
2170
ACTTTAACTGTAATGCATTTTAATGATAGCTATTCTACGTTTTACTATAGTTTATTAGTATCTAAGCCAG
T L T V M H F N D S Y S T F Y Y S L L V S K P
2240
TTTACAACTTTCATTTAAATGTTTCCAAAGTTAAAATTAATGAACATTGGTGCCAAGTAACAGCCACTTG
V Y N F H L N V S K V K I N E H W C Q V T A T C
2310
CGCACCTCCTGACAACCGCCCATACTTAAGTTACTTTAAATGGGAAGATCGCATTGAAAAAACAAACTAC
A P P D N R P Y L S Y F K W E D R I E K T N Y
2380
AAAGTTGTCATGCTTAATACTTCTGCTAACTCTAAACAGCCATTGTGCTGTCATCTTTATACACCAAGCT
K V V M L N T S A N S K Q P L C C H L Y T P S
2450
ACAACTATTATGATGAAATTGACATCTCAGAAGCGTGTATAGGCAGAGAAAAACTTTTACATAAAAATGA
Y N Y Y D E I D I S E A C I G R E K L L H K N E
2520
AAACTGGCTTCCTGTGCTGATTGTTGCTCCTATTTGTTTTTTAATTATTGGAATAATAATTATTTGTACT
N W L P V L I V A P I C F L I I G I I I I C T

```
        +         +         +         +         +         +      2590
TTTTGCAAGTTTAATTCTAAATACACGTTGTATGCTTTACTTATTCACAAAGGCGCCGATGCAGCCGAGT
 F  C  K  F  N  S  K  Y  T  L  Y  A  L  L  I  H  K  G  A  D  A  A  E
        +         +         +         +         +         +      2660
TAATGCTACTCACTATGACTCTAATTATCAAAACTTTATCACAATTTTGGCTGTATTAAGTGGAATTAGC
L  M  L  L  T  M  T  L  I  I  K  T  L  S  Q  F  W  L  Y  *

        +         +         +         +         +         +      2730
TTTTTTGTTGGTTTTGTTCACTTAATTTTATGGTGCTTGTTTAAATTTTACAAAAAATAAAATCTTACCT

        +         +         +         +         +         +      2800
AAGCTCTTTCACAAAATCGTCTTTGTGGATTTTTCGGGCTTCGATCCAAGCTCCTTCATGCAAACTCAAG

        +         +         +         +         +         +      2870
TACTTAATGTTAAGTTTAACGGCCACTTTCCTGAAACTTTTAAAAGACACTTTAAAGCAACCTAACTGCT

        +         +         +         +         +         +      2940
CACAGTTGTATTTTACTTCCATTCCAGCATGAAAGCAGCGTGCCAGAAGACTTTGATCCTGTTTACCTTA
                   ORF 14  M  K  A  A  C  Q  K  T  L  I  L  F  T  L
        +         +         +         +         +         +      3010
CGGAAAAAGACCACAAACATTCTGCCACCGTTTTATAGCTCAAATGGTTTTGTGGAGGCACCAACTGAAA
R  K  K  T  T  N  I  L  P  P  F  Y  S  S  N  G  F  V  E  A  P  T  E  T
        +         +         +         +         +         +      3080
CTCTTTCTTTAAAATTGGCCAATCCTGTTGATTTTATGCAAAATGGGGCTATTGGGTTAAAACTTGGAGG
  L  S  L  K  L  A  N  P  V  D  F  M  Q  N  G  A  I  G  L  K  L  G  G
        +         +         +         +         +         +      3150
TGGACTTTCAATTAACCAAGACGGCGAACTTGAATCTCAAACAATTACGTCTACAGTAAATCCACCTTTA
 G  L  S  I  N  Q  D  G  E  L  E  S  Q  T  I  T  S  T  V  N  P  P  L
        +         +         +         +         +         +      3220
TATCAACAAAATGGCGGTTTAAATTTAAAATATGGAGAAGAATTTGATATTGAAAATGAAGCTCTTAAAA
Y  Q  Q  N  G  G  L  N  L  K  Y  G  E  E  F  D  I  E  N  E  A  L  K  I
        +         +         +         +         +         +      3290
TAAAAACCATCGCTCCAATTACTAAGACAGAAAATGGATTAACTTTGTCTATAGGAGATGGATTAGAACT
  K  T  I  A  P  I  T  K  T  E  N  G  L  T  L  S  I  G  D  G  L  E  L
        +         +         +         +         +         +      3360
AAACTCAAATAATACACTCCAAGCTAGATTATCAAGTGGATTAGAAATTGACAACCAAGCAATTAGACTC
 N  S  N  N  T  L  Q  A  R  L  S  S  G  L  E  I  D  N  Q  A  I  R  L
        +         +         +         +         +         +      3430
AGAGTTCATGAACCTTTAAACTTAAATGCTTCTACTGGTGCTCTTCAATGCAGAATAGGAAATGGATTAA
R  V  H  E  P  L  N  L  N  A  S  T  G  A  L  Q  C  R  I  G  N  G  L  T
        +         +         +         +         +         +      3500
CAGTATCTGATAATAGCTTAGTTGTCTATCCTCATGAACCTTTAAATTTAGACCAAACTTCTGGCAAATT
  V  S  D  N  S  L  V  V  Y  P  H  E  P  L  N  L  D  Q  T  S  G  K  L
        +         +         +         +         +         +      3570
ACAACTAAGAGTGGGTAATGGAATTAATGTGCAAAATAGCTCGCTTGTTGCAAGAATAGGGCAGGGATTG
 Q  L  R  V  G  N  G  I  N  V  Q  N  S  S  L  V  A  R  I  G  Q  G  L
        +         +         +         +         +         +      3640
GCTTTTAACAACTCTGACATTCAAATTAATGCCGCACCTCCATTTACTTTTTCTAATAATCAGCTTTCTA
A  F  N  N  S  D  I  Q  I  N  A  A  P  P  F  T  F  S  N  N  Q  L  S  I
        +         +         +         +         +         +      3710
TAGCACTAGGGGATGGATTAGTAACTAATGCAAGTCAATTAAAAGTTAATTTTGGTAAAGGACTTTTTAT
  A  L  G  D  G  L  V  T  N  A  S  Q  L  K  V  N  F  G  K  G  L  F  I
        +         +         +         +         +         +      3780
TAATTCTTCTGATTCATCAAAACTTCAAGTAAACATTAGACCTCCTTTAAATTATTTTGGAAATAGCAAT
 N  S  S  D  S  S  K  L  Q  V  N  I  R  P  P  L  N  Y  F  G  N  S  N
```

```
        +         +         +         +         +         +      3850
TCTCTTACTGTAGTTACTGGTAATGGCTTAGGAGTATCTGGTACTAATCTAGGCTCTAACCTTTATGTTA
S  L  T  V  V  T  G  N  G  L  G  V  S  G  T  N  L  G  S  N  L  Y  V  K
        +         +         +         +         +         +      3920
AAACTGGCAATGGATTAGAAGCTGATAGCAGTAACGTTAGAGTAAAAATTGCAAATGGGTTACAGTTTAC
  T  G  N  G  L  E  A  D  S  S  N  V  R  V  K  I  A  N  G  L  Q  F  T
        +         +         +         +         +         +      3990
TGATGGCAACATTGAAGCTAATTTAGGAAATGGGTTAACATTTTCAAACGGTCAAATTACTGCAAACATT
 D  G  N  I  E  A  N  L  G  N  G  L  T  F  S  N  G  Q  I  T  A  N  I
        +         +         +         +         +         +      4060
GGCGCCGGTCTTGCGTTTTTAAATGGTCAAATTACACTAGTGAACAGCACTCCTTCTGGTTATACAGATT
G  A  G  L  A  F  L  N  G  Q  I  T  L  V  N  S  T  P  S  G  Y  T  D  Y
        +         +         +         +         +         +      4130
ATACTTTGTGGACTACTCCAGACCCTTCTCCTAATGCTAGCATTAAAACTGACTTAGATGCTAAGCTTGT
  T  L  W  T  T  P  D  P  S  P  N  A  S  I  K  T  D  L  D  A  K  L  V
        +         +         +         +         +         +      4200
TTTAACTTTATCAAAAGCAGGCAGCACTGTAATAGGCACAATAGGTATTTTTGCTCTAAAATCTCCTCTT
 L  T  L  S  K  A  G  S  T  V  I  G  T  I  G  I  F  A  L  K  S  P  L
        +         +         +         +         +         +      4270
ACTCCTATTTCAGAAAATTCTATTAATGTTGAAATTTTTTTTGATGCTAATGGAGAAATTAATTTAACTA
T  P  I  S  E  N  S  I  N  V  E  I  F  F  D  A  N  G  E  I  N  L  T  T
        +         +         +         +         +         +      4340
CTAGTTCGCTAAAAAGCTATTGGGGATTTAGAGAAGGTGATTCTTATAATCCATCTTCTAACTTAAACCC
  S  S  L  K  S  Y  W  G  F  R  E  G  D  S  Y  N  P  S  S  N  L  N  P
        +         +         +         +         +         +      4410
CCTTTACTTGATGCCTAATACTTATGCATACCCTCAAGGTCGGAAAACTATTACACAAGTTTTTCCACTA
 L  Y  L  M  P  N  T  Y  A  Y  P  Q  G  R  K  T  I  T  Q  V  F  P  L
        +         +         +         +         +         +      4480
GAAGTATACTTAAATGGAGACACTGCTAAGCCTGTGCCCTTAGAAGTTGCTTTTAATACTTTGTCTTCTA
E  V  Y  L  N  G  D  T  A  K  P  V  P  L  E  V  A  F  N  T  L  S  S  T
        +         +         +         +         +         +      4550
CTGGTTTTTCTCTTGAATTTACTTGGAGAAATCTTAATGCTTACACTGGAGAAGCGTTTCGTGTCTTTAG
  G  F  S  L  E  F  T  W  R  N  L  N  A  Y  T  G  E  A  F  R  V  F  R
        +         +         +         +         +         +      4620
GAAATTTTACATACATTAGTCAATATTAAATTTTAAAACTTTTTATTGCTGATTTTGGTAATACACGACA
 K  F  Y  I  H  *
        +         +         +         +         +         +      4690
TGTTAACATTCCACCACCTTCCCATTTTACTTTGTAAACAAATCTATTAAAATGCAATCCAGTGTTGTAA

        +         +         +         +         +         +      4760
TTATCTTGATTGTGTCTTAATTTTACAGCATGTTCTACACATTCAAAATCAGGGCTAGTAATCAACACAA

        +         +         +         +         +         +      4830
ATCCGGTAAAGCTAGAGCTTAAATCAGGCACACAGTTTAATAATTGGGGTCGTAGTAAATGTCGTGCTTC
                                                ORF 17    M  S  C  F
        +         +         +         +         +         +      4900
GGACGCTTGTGCCTTCCTCGTTCTGCGTCAAAAAGCACTGTGGACATCCGCAGCTGTGTTTCACCATCTC
G  R  L  C  L  P  R  S  A  S  K  S  T  V  D  I  R  S  C  V  S  P  S  P
        +         +         +         +         +         +      4970
CACCTTGCAAAGCGGGGTGCCCTTCGGATCACGAGCTCCGCACGTCTAGCGGTGAGTTCAGCTTCAGCTC
  P  C  K  A  G  C  P  S  D  H  E  L  R  T  S  S  G  E  F  S  F  S  S
        +         +         +         +         +         +      5040
TAGCACAGGCTTCGTAAAAGCGTCCCATACCTTGCACCCAGAAGATTTTAAATCACCATCTAACCGCTTA

 S  T  G  F  V  K  A  S  H  T  L  H  P  E  D  F  K  S  P  S  N  R  L
        +         +         +         +         +         +      5110
GCTCAGAAACACTACATTCAGCATGCCAGCCAAAAGGTAAATCAAATTCTACATCAAAGAAAGCGGTAGC
A  Q  K  H  Y  I  Q  H  A  S  Q  K  V  N  Q  I  L  H  Q  R  K  R  *
```

```
         +         +         +         +         +         +      5180
AGAGCCAACTTTTCTTACTACTATATCTTTCTCTGTGCGTAAATGGGCAATATCTTCATTATTTTTATCT

         +         +         +         +         +         +      5250
CTAATAGGAGCAAACTTTGTGCTTTGATACATGCACCTTACACGTTTGTTCCTATTTTGCCAAGCTAGTC

         +         +         +         +         +         +      5320
TTTCAGCTTCCGTCATTTTTTCCCATCTACGTTCAAAGCGTTTCTGATTTCTATCCTCACATTTCTTATG

         +         +         +         +         +         +  5386
TCCACAATCATCAAACTCACAAGAAAAGTATCTATATCATTAAAATTGAGGTTCCAGCTGGAATTC
```

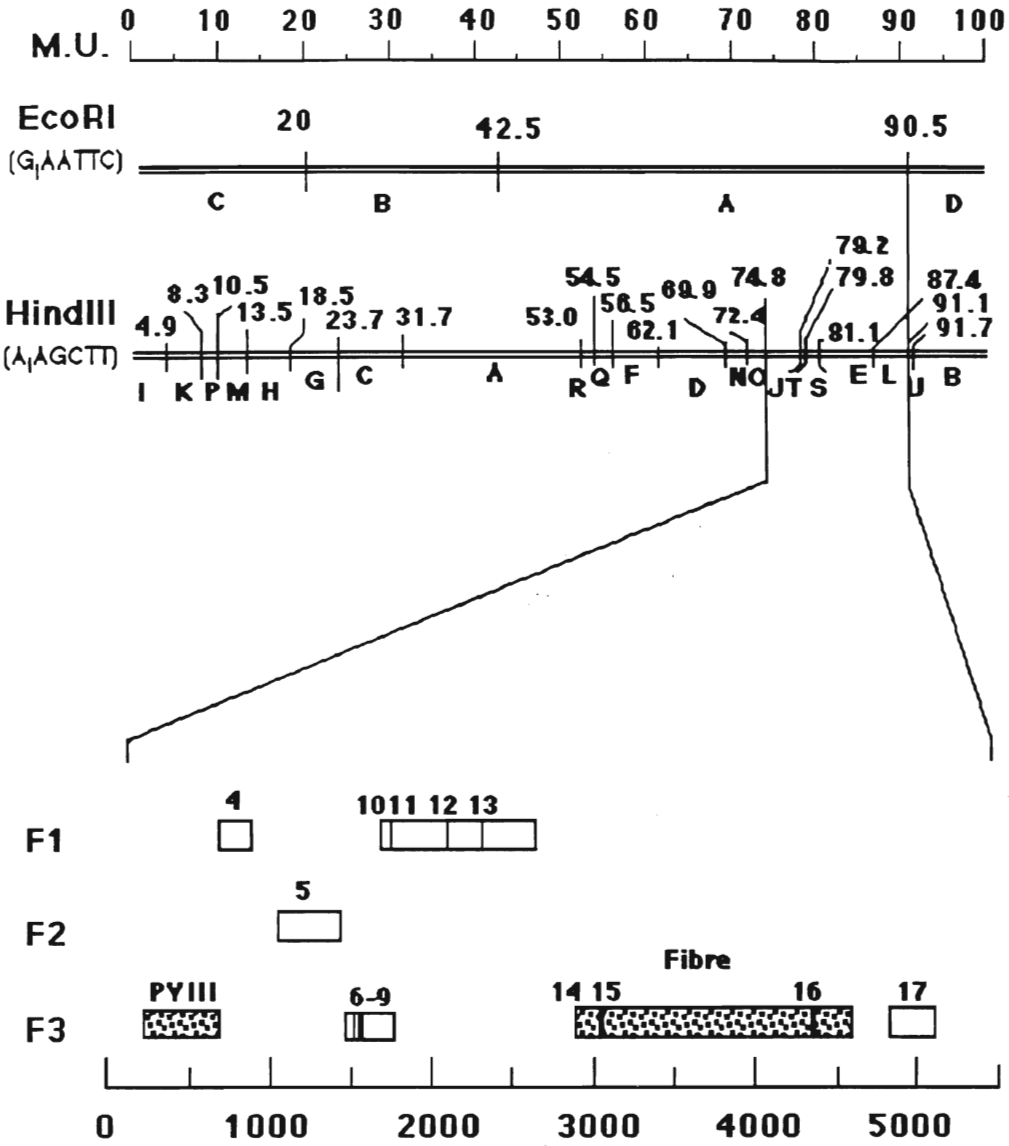


Figure 13:

The genome of BAV2 showing the location of the *Eco*RI and *Hin*dIII sites and the deduced open reading frames of the 5375 bp fragment from 74.8 to 90.5 mu. ORFs for the r strand which can encode 50 amino acids are indicated and numbered sequentially. The two shaded ORFs indicate regions with identity to pVIII and fibre proteins of known adenovirus sequences

Table 1

Predictions on coding potential of open reading frames in BAV2 putative E3 and neighbouring sequences determined using the TESTCODE algorithm (Fickett, 1982).

ORF #	ORF	Probability of coding	Prediction
1	pVIII - 16.8 KDa (ATG_{225})	0.77	No Opinion
4	E3 - 7.1 KDa(ATG_{685})	0.92	Coding
5	E3 - 14.6 KDa(ATG_{1052})	0.29	Noncoding
6	E3 - 11.7 KDa(ATG_{1416})	0.92	Coding
10	E3 - 37.1 KDa (ATG_{2319})	0.77	No Opinion
14	Fibre - 59.6 KDa (ATG_{2898})	0.92	Coding
17	? - (ATG_{4358})	0.04	Noncoding

only ORFs 1 and 14 showed substantial homology with any known adenoviral sequence.

4.3.2 Corrected BAV2 genome *Hin*dIII Map

The nucleotide sequence analysis revealed the existence of 2 additional *Hin*dIII sites. The sizes of these fragments could not have been determined by restriction endonuclease analysis and agarose gel electrophoresis. The 183 and 437 bp fragments were located between 79.2 and 81.1 mu. They were placed in the physical map of pEA (fig. 6) and the original linear map of BAV2 (constructed in collaboration with Kirsty Salmon, Brock University, St.Catharines, Ontario, Canada) was modified (fig. 10, 11, 13).

4.3.3 Organization and homology to the pVIII region

ORF 1 encodes a polypeptide of 156 amino acids. Comparisons with published protein sequences of hexon associated protein precursor pVIII revealed an identity of 53, 54 and 33 percent with Ad41, Ad 5 and MAV1 respectively. Figure 14 shows a comparison between the predicted amino acid sequences of pVIII of BAV2 and Ad2. An identity of 54 percent and a similarity of 82 percent of the amino acid sequences was seen. From alignment, it was possible to recognize that the C- terminus was highly conserved. An identity of 63 percent and a similarity of 87 percent was determined for the last 70 amino acids.

This region contained the TATA box and the 5' E3 cap site of mRNA found in human adenoviruses. The characteristic sequence TATAA and cap site were found in BAV2 at nucleotide 374 and 29 base pairs away at nucleotide 403 respectively. The deduced amino acid sequence of ORF 1 showed the greatest homology to the region that corresponded to the Ad2 E3 leader sequence. The

```
                                                  10        20        30
BAV2                                            MNPPAWPASRLFQPVDTVQTLELPRNELLE
                                                :||::|||: ::|   :  |: |||::  |
Ad2     HMISRVNGIRAHRNRILLEQAAITTTPRNNLNPRSWPAALVYQESPAPTTVVLPRDAQAE
        40         50         60         70         80         90


                 40              50         60         70          80
BAV2    TAMTNSGMQLAGGGVHRTK----DIKPEDLVGRGIQLNSYQPPTT-RLRPDGVFQLAGGS
        ::|||||| |||||  ||::    :|:: :: ||||||||:   :::  |||||:||::|::
Ad2     VQMTNSGAQLAGGFRHRVRSPGQGITHLKIRGRGIQLNDESVSSSLGLRPDGTFQIGGAG
       100        110        120        130        140        150


          90        100        110        120        130        140
BAV2    RSSFNPSINTLLTLQPAASVPRSGGIGEVQFVHEFVPSVYFQPFSGPPGSYPDEFIYNFD
        ||||:|: :::|||||:::| |||||||::||::|||||||:|||||| |||:|| |||
Ad2     RSSFTPR-QAILTLQTSSSEPRSGGIGTLQFIEEFVPSVYFNPFSGPPGHYPDQFIPNFD
       160        170        180        190        200        210


          150
BAV2    VATDSIDGYAZ
        :::|| |||
Ad2     AVKDSADGYD
        220
```

Figure 14:

A comparison between the predicted amino acid sequence of ORF 1 of BAV2 and the Ad2 hexon associated pVIII precursor using the FASTA program. Identical residues are indicated by a bar and similiar residues by a colon.

316 bp nucleotide sequence of the Ad2 leader was compared to the corresponding sequence of BAV2 (fig. 15). The identity was 81 percent. Also, putative the E3 leader sequence in BAV2 was the only one to show significant homology to BAV3 (fig. 16). There was 64 percent identity between BAV2 and BAV3 in the 285 bp overlap.

4.3.4 The organization and protein coding capacity of the putative E3 region in BAV2

The distance between the cap site and the most distal polyadenylation site before the start of the ORF that may code for the fibre gene is 2317 bp . The E3 regionpresented here is located between nucleotides 403 and 2720. ORFs were examined for similarity to known adenovirus proteins. ORFs 4-13 showed no substantial homology with any published adenoviral (EMBL) sequence E3 or otherwise. Only ORFs 4 and 6 were predicted using the TESTCODE algorithm to code for proteins. Potential N-glycosylation sites were located in all the putative E3 ORFs except ORF 4. ORFs 5 and 6 have two and ORF 10 has ten potential N -glycosylation sites.

The characteristics of ORF 10 indicate that it codes for a glycoprotein of 324 amino acids. The deduced amino acid sequence of the 37.1 KDa polypeptide contained a strongly hydrophobic sequence near the C-terminus with ten N-glycosylation sites between. The hydropathy profiles of the predicted amino acid sequence of this ORF and Ad5 gp19 KDa were compared (fig. 17). The long relatively hydrophilic region and the hydrophobic region near the C-terminus was similiar in both profiles.

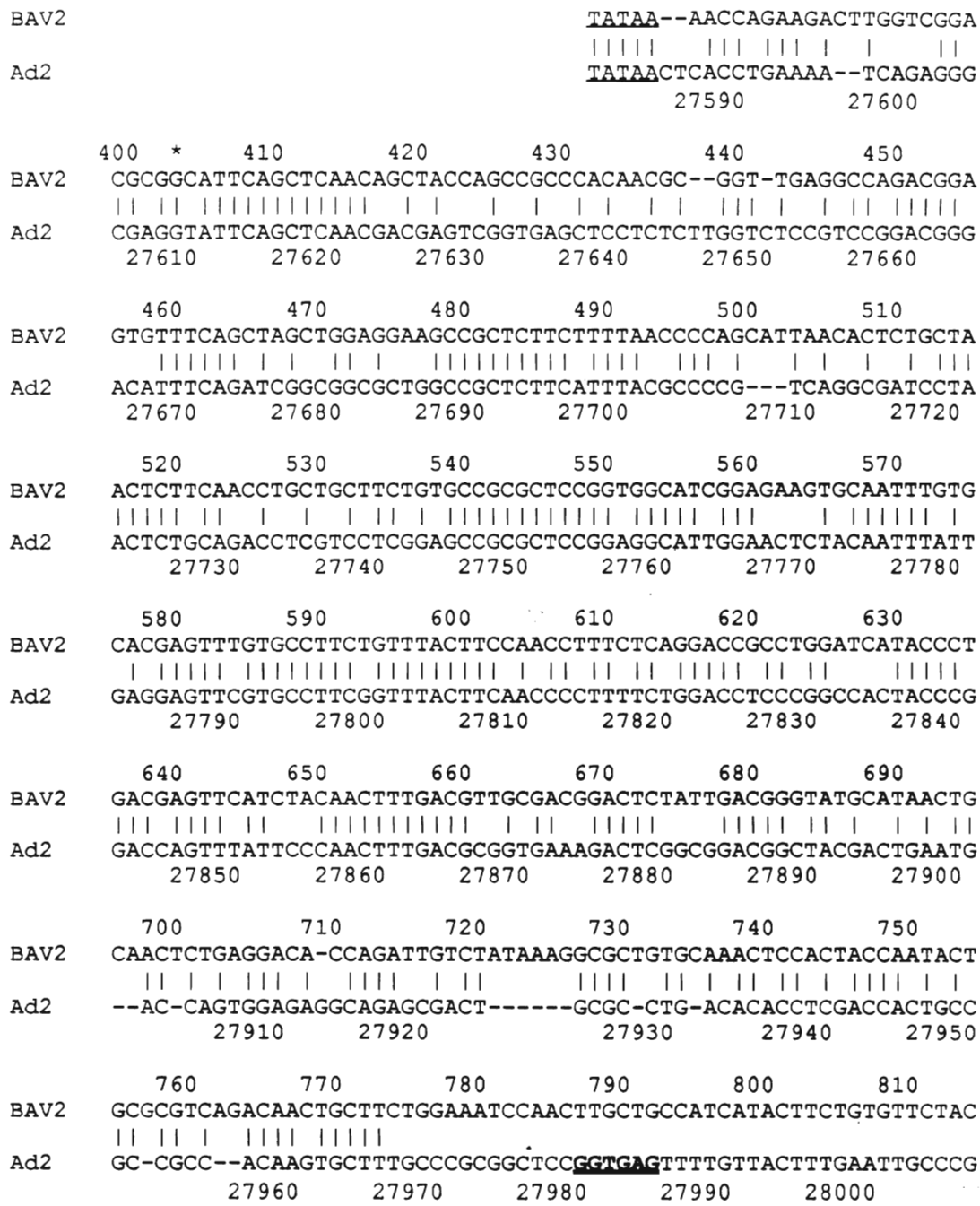


Figure 15:

Alignment of the Ad 2 E3 leader (lower) nucleotide sequence with BAV2 (upper) using the FASTA program (Pearson and Lipman, 1988). The identity is 66 percent. The asterisk indicates the start of the E3 region in Ad2. The TATA box and the end of the Ad2 leader sequence (the splice site) are underlined.

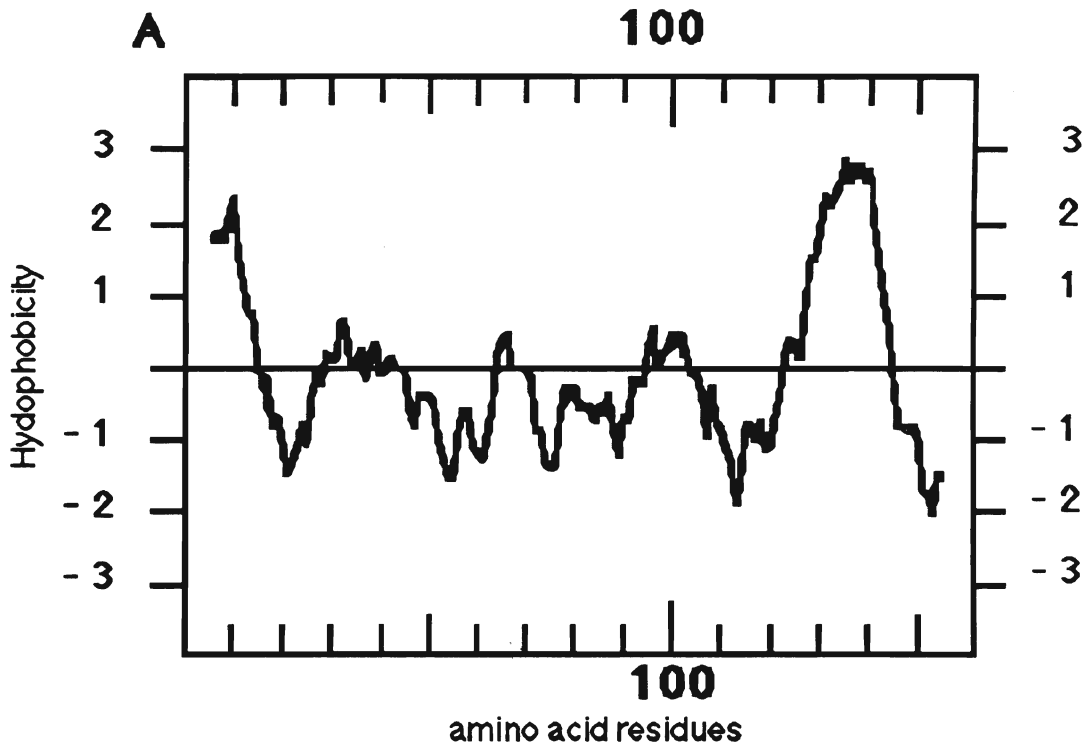
```
          380        390        400        410        420        430
BAV2   TAAAACCAGAAGACTTGGTCGGACGCGGCATTCAGCTC-AACAGCTA-CCAGCCGCCCAC
                                           |||||||| |||    || || ||    || |
BAV3   CGCTGCGGGCGATTACTTTAAAAGCCCCACTTCAGCTCGAACCCTTATCCCGCTCACCGC
         110        120        130        140        150        160

            440        450        460        470        480        490
BAV2   AACGCGGTTGAGGCCAGACGGAGTGTTTCAGCTAGCTGGAGGAAGCCGCTCTTCTTTTAA
         |  | || || ||||| ||||| ||||| |||| |||||    || || ||||| ||
BAV3   CTCCTGCTTAAGACCAGATGGAGTCTTTCAACTAGGAGGAGGCTCGCGTTCATCTTTCAA
         170        180        190        200        210        220

             500        510        520        530        540        550
BAV2   -CCCCAGCATTAACACTCTGCTAACTCTTCAACCTGCTG-CTTCTGTGCCGCGCTCCGGT
        |||| |||    | | | |||     |||| ||    ||| | ||    ||||||  |||
BAV3   CCCCCTGCAAACAGATTTTGC-----CTTCCACGCCCTGCCCTCCAGACCGCGCCACGGG
         230        240             250        260        270        280

              560        570        580        590        600        610
BAV2   GGCATCGGAGAAGTGCAATTTGTGCACGAGTTTGTGCCTTCTGTTTACTTCCAACCTTTC
       |||||  |||      ||| ||||| | || ||||||||  | || ||| || | || | |
BAV3   GGCATAGGATCCAGGCAGTTTGTAGAGGAATTTGTGCCCGCCGTCTACCTCAACCCCTAC
              290        300        310        320        330        340

              620        630        640        650        660        670
BAV2   TCAGGACCGCCTGGATCATACCCTGACGAGTTCATCTACAACTTTGACGTTGCGACGGAC
       || |||||||| |  || || || ||| |||| ||   | |||   ||||    |   ||
BAV3   TCGGGACCGCCGGACTCTTATCCGGACCAGTTTATACGCCACTACAACGTGTACAGCAAC
              350        360        370        380        390        400

              680        690        700        710        720        730
BAV2   TCTATTGACGGGTATGCATAACTGCAACTCTGAGGACACCAGATTGTCTATAAAGGCGCT
       ||| |   ||| |||
BAV3   TCTGTGAGCGGTTATAGCTGAGATTGTAAGACTCTCCTATCTGTCTCTGTGCTGCTTTTC
              410        420        430        440        450        460
```

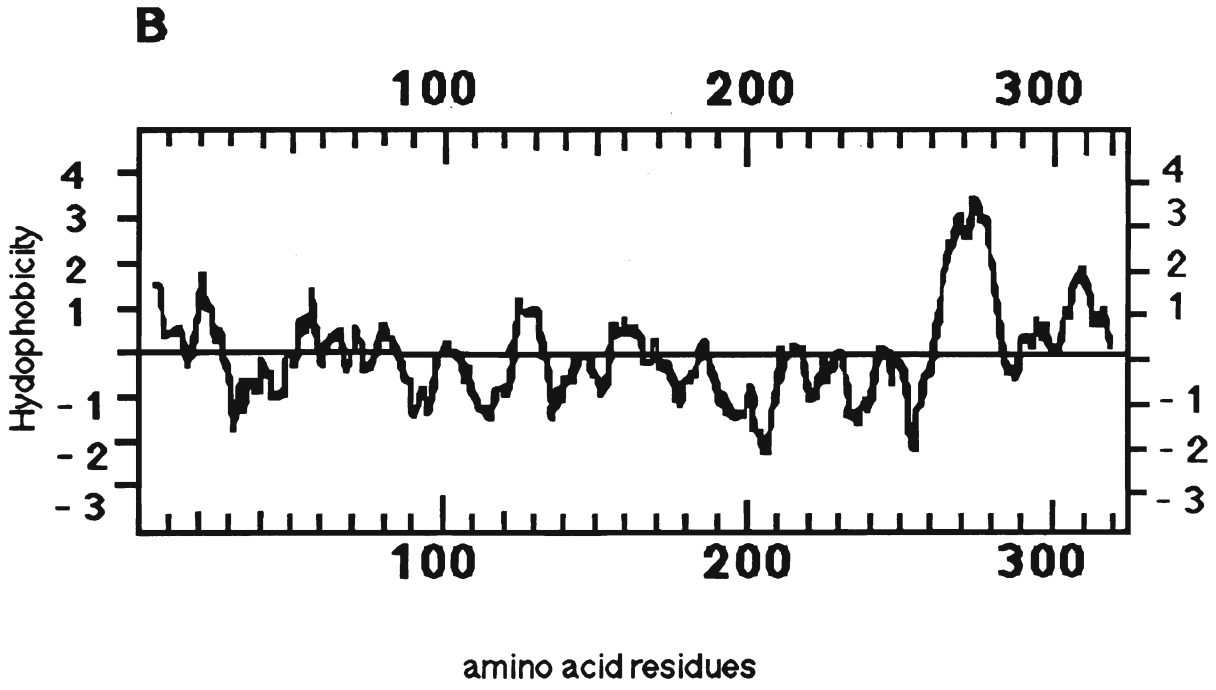
Figure 16:

Alignment of the BAV3 nucleotide sequence with BAV2 (upper) using the FASTA program (Pearson and Lipman, 1988). The identity is 64 percent in the 285 bp overlap. This nucleotide sequence corrresponds to the Ad2 leader sequence and pVIII.

Figure 17:

Hydropathic profiles of Ad5 gp19 kDa (A) and the putative BAV2 E3 protein 37.1 KDa (B) using the Kyte-Doolittle (1982) hydropathy scale. The protein patterns are drawn to the same scale. The hydrophobicity is displayed above the abscissa and the hydrophilicity below.


A
100
3
2
1
- 1
- 2
- 3
Hydophobicity
100
amino acid residues


B
100
200
300
4
3
2
1
- 1
- 2
- 3
Hydophobicity
amino acid residues

4.3.5 Organization and homology of the fibre gene region

The proposed BAV2 fibre gene (ORF 14) is encoded starting at nucleotide 2898 and ending at 4568. ORF 14 encodes a polypeptide of 557 amino acid and a theoretical molecular weight of 59.6 KDa. Nine potential sites for N-glycosylation were identified (fig. 12). Potential N-glycosylation sites in the deduced amino acid sequence of the fibre gene have been previously reported (Raviprakash *et al.*, 1989; Hong *et al.*, 1988; Mei & Wadell, 1993). Comparisons with published protein sequences of the fibre gene revealed an identity of 30.2, 27.5 and 24.6 percent to Ad5, Ad2 and Ad40 respectively. Figure 18 shows a comparison between the predicted amino acid sequences of the fibre polypeptides of BAV2 and Ad41. There was an identity of 27.7 percent and asimilarity of 64 percent between the amino acid sequences. Figure 19 compares the translated sequences of Ad41 and BAV2 arranged in the three regions of the adenoviral fibre -tail, shaft and knob, according to Green *et al.* (1983). The similarity found between Ad41 and BAV2 was greatest in the shaft region. According to the alignment determined by the FASTA program, the tail region in BAV2 was 36 residues and had an identity of 28 percent and a similarity of 42 percent with Ad41. The shaft region was predicted to lie between residues 37 and 390 and had an identity of 28 percent and a similarity of 73 percent with Ad 41. The knob, which was 166 amino acids long, started at residue 391 had an identity of 32 percent and a similarity of 65 percent with Ad41. The most significant homology was between residues 390 and 410, the transition between the shaft and the knob. This 20 amino acid region in BAV2 had an identity of 70 percent and a similarity of 95 percent with Ad40 and Ad41. Also, these 20 residues in BAV2 have an identity of 70 percent and 60 percent and a similarity of 85 percent and 90 percent with Ad2 and Ad5 respectively.

Figure 18:

Comparison between the predicted amino acid sequence of ORF14 of BAV2 and the Ad41 fibre protein using the FASTA prgram (Pearson and Lipman, 1988). Identical residues are indicated by a bar and conserved residues by a colon.

```
                  10        20        30        40        50
BAV2        MKAACQKTLILFTLRKKTTNILPPFYSSNGFVEAPTETLSLKLANPVDFMQNGA
                        :  :: | ||| |||| : | |:::|||| ::|:: :|||
Ad41   MKRARLEDDFNPVYPYEHYNPLDIPFITPPFASSNGLQEKPPGVLSLKYTDPLT-TKNGA
              10        20        30        40        50

          60        70        80        90       100       110
BAV2   IGLKLGGGLSINQDGELESQTITSTVNPPLYQQNGGLNLKYGEEFDIENEALKIKTIAPI
       ::|||||:||:|:::|:|:|:: : :|:::|: : |  ::|:|:: ::: ::||::: ||:
Ad41   LTLKLGTGLNIDENGDLSSDA-SVEVSAPITKTNKIVGLNYTKPLALRSNALTLSYNAPL
      60        70        80         90       100       110

         120       130       140       150       160       170
BAV2   TKTENGLTLSIGDGLELNSNNTLQARLSSGLEIDNQAIRLRVHEPLNLNASTGALQCRIG
       : ::|:|:|:|:: :::|:|||:|   ::: |: |: ::|||:  :|||:| ::|  |:  ::
Ad41   NVVNNNLALNISQPVTVNANNELSLLIDAPLNADTGTLRLQSAAPLGLVDKT--LKVLFS
      120       130       140       150       160       170

         180       190       200       210       220       230
BAV2   NGLTVSDNSLVVYPHEPLNLDQTSGKLQLRVGNGINVQNSSLVARIGQGLAFNNSDIQIN
       : | :::| |::  : ||:|: ::: : |: :  ::::|::|: ::|  :: :::::::::
Ad41   SPLYLDNNFLTLAIERPLALSSSRA-VTLKYSPPLKIENENLTLSTGGPFTVSGGNLNLT
        180       190       200       210       220       230

         240       250       260       270       280       290
BAV2   AAPPFTFSNNQLSIALGDGLVTNASQLKVNFGKGLFINSSDSSKLQVNIRPPLNYFGNSN
       :::|::  || ||:::::: | : :| | |: :  : |::|:   |::::   |:   :::
Ad41   TSAPLSVQNNSLSLVITSPLKVINSMLAVGVNPPFTITDSG---LAMDLGDGLAL--GGS
         240       250       260       270          280          290

         300       310       320       330       340           350
BAV2   SLTVVTGNGLGVSGTNLGSNLYVKTGNGLEADSSNVRVKIANGLQFTDG----NIEANLG
       :|::  | || :|:::::: :| :       ::  ::|:  ::|| :::     |::|| |
Ad41   KLIINLGPGLQMSNGAITLALDAALPLQYRDNQLQLRIGSTSGLIMSGVTQTLNVNANTG
              300       310       320       330       340       350

              360       370       380       390       400       410
BAV2   NGLTFSNGQITANIGAGLAFLNGQITLVNSTPSGYTDYTLWTTPDPSPNASIKTDLDAKL
       :||: :|: :::::|:|| |  :  : :: :|:: |  |||||:||||||:: ::|||||:
Ad41   KGLAVENNSLVVKLGNGLRF--DSWGSITVSPTTTTPTTLWTTADPSPNATFYESLDAKV
              360       370         380       390       400

              420       430       440       450               460
BAV2   VLTLSKAGSTVIGTIGIFALKSPLTPISENSINVEIFFDANGE--------INLTTSSLK
        |:| | :: | |||:| | |: | : ::: |:  ::| ::|:         | :: : :
Ad41   WLVLVKCNGMVNGTISIKAQKGILLRPTASFISFVMYFYSDGTWRKNYPVFDNEGILANS
      410       420       430       440       450       460
```

```
            470        480        490        500        510        520
BAV2   SYWGFREGDSYNPSSNLNPLYLMPNTYAYPQGRKTITQVFPLE-VYLNGDTAKPVPLEVA
       : ||:|:|:| |:: : |:: :||::  ||: : : :| ::|: ::|:||::::::::
Ad41   ATWGYRQGQSANTNVS-NAVEFMPSSKRYPNQKGSEVQNMALTYTFLQGDPNMAISFQSI
      470        480        490        500        510        520

            530        540        550
BAV2   FNTLSSTGFSLEFTWRNLNAYTGEAFRVFRKFYIHZ
       :|  : :|:||:|||| |    :      :: :
Ad41   YNH-ALEGYSLKFTWRVRNNERFDIPCCSFSYVTEQ
       530        540        550        560
```

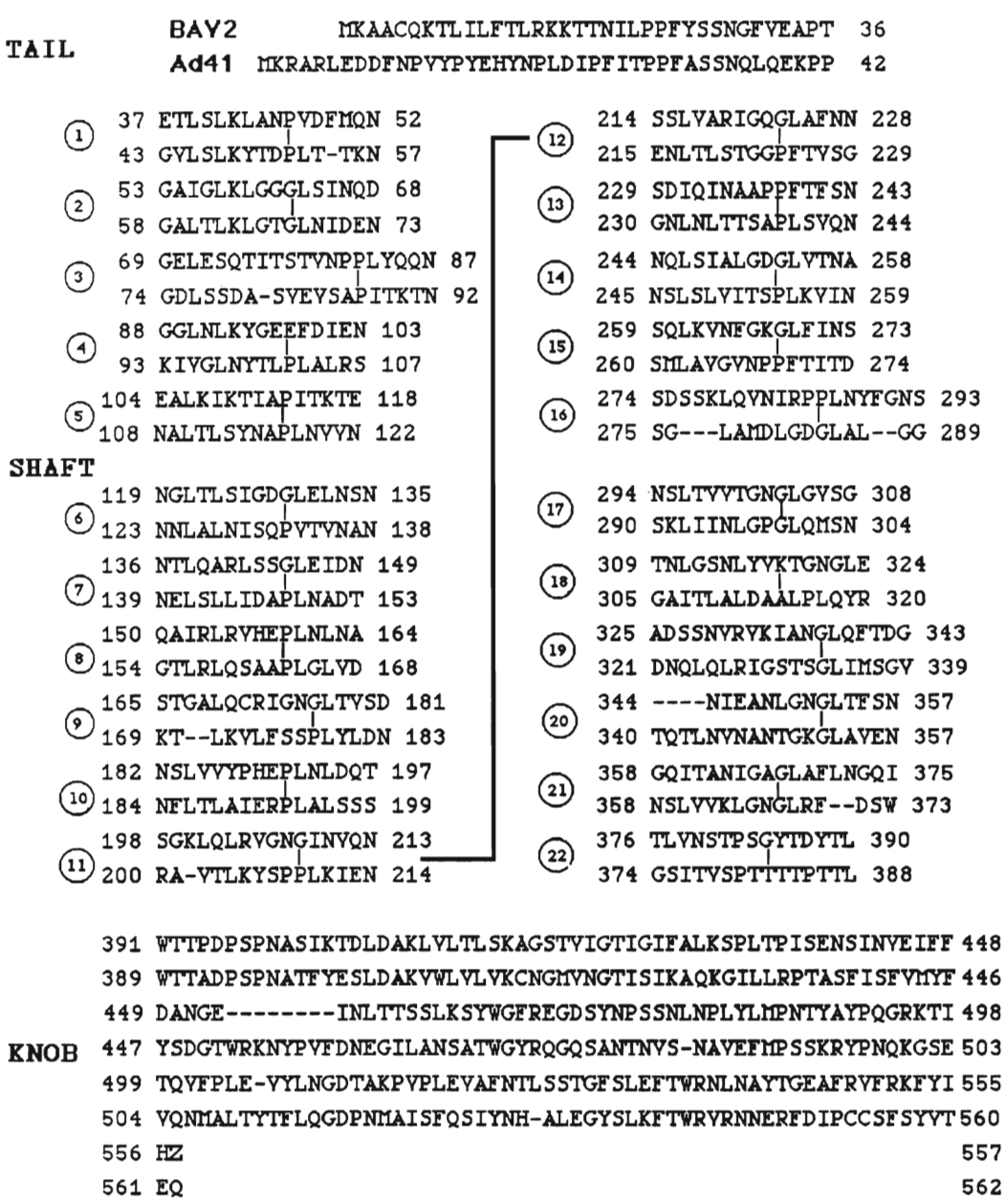


Figure 19:

Organization of the subunits of Ad41 and BAV2 fibre proteins. Each fibre protein has a short tail, separated from a distal knob by a repetitive sequence which forms the shaft. The part of the figure which illustrates the organization of the Ad41 fibre is taken from Kidd *et al.* (1990). Position a5, usually a proline (P) or glycine (G), is indicated with a bar.

The basic structural unit of the model proposed by Green *et al.* (1983) for the shaft is made up of 15 amino acids (fig. 20 A). Twenty two, 15 residue motifs were found from amino acid 37 to 390 in BAV2 ORF 14. In the arrangement shown in figures 19 and 20 B it was noted that there was a periodicity of hydrophobic residues . There are 108 hydrophobic residues in the proposed shaft of BAV2; 67 are not conserved, but 45 are replaced by another hydrophobic residue. In position a5 there is nearly always a proline or a glycine. Following the residue numbering of Ad41 proposed by Kidd *et al.* (1990) there were 13 glycines and 7 prolines in position a5. A glutamic acid and a lysine were found in position a5 in motif 4 and 18 respectively.

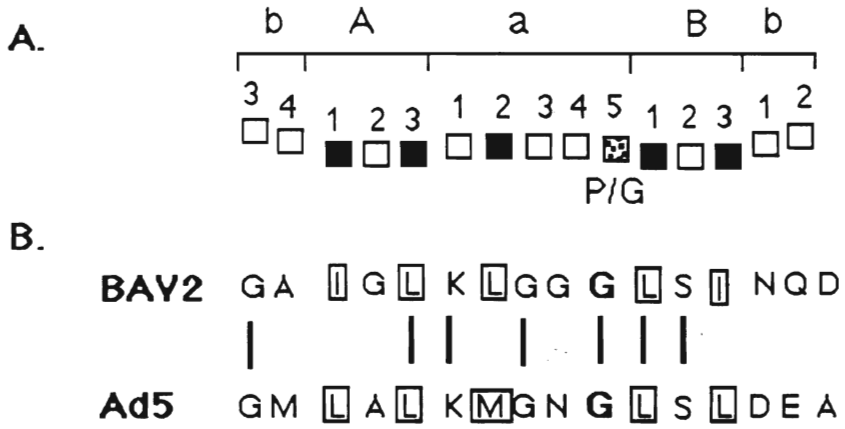


Figure 20:

A: Structure of the fibre, a repeating motif sequence based on the amino acid sequences in the shaft regions of Ad2, Ad5, Ad3, Ad40 and Ad41 fibres. Position a5 is almost always a proline or a glycine, although occasionally a lysine residue has been found in this position. The closed boxes are occupied by hydrophobic residues. **B:** Comparison of the deduced amino acid sequence between BAV2 and Ad5 of the proposed motif 2. The identical residues are indicted with a bar. The hydrophobic amino acids in the motif are shown in boxes. The glycine in position a5 is boldfaced.

Chapter 5: Discussion

The nucleotide sequence of BAV2 between 74.8 and 90.5 mu was determined. Two of the17 ORFs which could encode polypeptides of 50 amino acids or more (ORF 1 and ORF 14) exhibited homology to the well characterized protein pVIII and the fibre protein of human adenoviruses. ORF1, located between 75.5 and 76.9 mu, encoded a polypeptide of 156 amino acids and showed homology to protein pVIII. The deduced amino acid sequence from ORF14 that lies between 83.7 and 88.9 on the BAV2 genome showed homology to the fibre protein. The nucleotide sequences or the deduced amino acid sequences of the remaining 15 ORFs (12 of which lie between ORF1 and ORF14) showed no substantial homology to any known adenoviral DNA or protein sequences.

5.1Organization and homology to the pVIII region

Comparisons with Ad2, Ad5, Ad41, Ad3 and MAV1 using the EMBL database suggest that ORF 1 encodes the hexon associated pVIII gene products. The amino acid sequence predicted from BAV2 ORF 1 showed an overall identity of about 54 percent with human adenoviruses and 33 percent identity to MAV1. Although the identity of this BAV2 amino acid sequence is 54 percent with Ad2, the similarity of 82 percent with Ad2 is significant. This is not surprising because protein VIII is highly conserved among adenovirus serotypes

(Cladaras & Wold, 1985; Raviprakash *et al*., 1989; Dragulev *et al.*, 1991; Mittal *et al.*, 1992). ORF 1 can encode a polypeptide of 156 amino acids. The N-terminus pVIII is absent in BAV2 when compared to the 227 amino acid sequence of Ad2 and Ad5 and the 215 amino acid sequence of MAV1 (Raviprakash *et al.*, 1989). The significance of this is unknown, however it may be that the BAV2 ORF1 undergoes complex splicing to generate a similiar sized polypeptide.

The start of the Ad2 E3 promoter lies within the sequences encoding pVIII, about 320 bp downstream of the termination codon. The amino acid and nucleotide sequences of BAV2, compared to the 316 bp Ad2 leader sequence were highly conserved. The last 70 residues of BAV2 ORF 1 that lie within the E3 promoter region appeared to be highly conserved, showing an identity of 63 percent and a similarity of 87 percent with Ad2 pVIII. BAV2 shows 81 percent nucleotide homology to the 316 bp E3 leader sequence of Ad2. The consensus TATA box is found in a corresponding region in BAV2 The position of the TATA box and the nucleotide similarity with Ad2, suggest that this region corresponds to the E3 promoter in BAV2. The conserved capped G nucleotide at residue 403 in BAV2 locates the beginning of the E3 region at 76.1 mu.

5.2 Organization and protein coding capacity of the putative E3 region of BAV2

Three polyadenylation (poly A) signals were located at nucleotides 1652, 1682 and 2715. The E3 region in human adenoviruses is known to be located between the L4 pVIII gene and the L5 fibre gene (Cladaras & Wold, 1985). If this genomic organization is conserved then the poly A signal for the BAV2 E3 region transcripts would be at nucleotide 2715, 181 bp away from the beginning of ORF 14. Therefore, the BAV2 E3 region located between 76.1 and 83.3 mu

appears to be approximately 2.3 kbp in size. The putative BAV2 E3 region represents 7.1 percent of the viral genome compared to 9 percent in Ad5 (Cladaras & Wold, 1985). Interestingly, although the Putative BAV2 E3 region is smaller than Human Ads (Cladaras & Wold, 1985), it is almost twice the size of MAV1 (Raviprakish *et al.,* 1989) , CAV1 (Dragulev *et al.,* 1991) and CAV2 (Linne, 1992). Also, even though the BAV3 genome is larger than the BAV2 genome, the predicted 1.5 kbp BAV3 E3 region is smaller than that of BAV2 and only represents 4.3 percent of the viral genome (Mittal *et al.*, 1992). The significance of this is unknown, however it may be related to the pathogenicity of these two viruses. Both are known to cause respiratory and enteric diseases in cattle (Mohanty, 1971; Belak & Palfi, 1974; Lehmkuhl *et al.,* 1975). However, experimental results produced by the subgroup 1 bovine adenoviruses show that BAV3 is the most virulent (reviewed by Lehmkuhl & Gough, 1977). It may be that the larger E3 region in BAV2 functions to modulate the host response.

The data suggested that BAV2 E3 region, while it is transcribed rightward from approximately the same map coordinate, differed significantly from the E3 of human adenoviruses. The E3 regions of Ad2 and Ad5 are located between the L4 pVIII and the L5 fibre gene and encode at least 10 polypeptides (Cladaras & Wold, 1985). These polypeptides can be divided into two families depending on the polyadenylation signal used. Twelve ORFs in BAV2, having the potential to encode polypeptides larger than 50 amino acids after an initiation codon, were located between the pVIII and the fibre gene. The largest ORFs (4, 5, 6, 10) can encode polypeptides of theoretical molecular weights 7.1, 14.6, 11.7, and 37.1 KDa, respectively. The 12 ORFs in BAV2 exhibit no significant homology to any known adenoviral DNA or protein sequence. Two of the 4 ORFs were predicted using the TESTCODE algorithm to code for proteins (Fickett, 1982). ORF 5 is predicted to be noncoding. It is unlikely that ORF 10 is noncoding because it has

similar characteristics to the Ad5 gp19 KDa. The well characterized E3 gp19 KDa has been found to be one of the most conserved E3 proteins (Paabo *et al.,* 1986). The E3 gp19 KDa is found to exist in human adenoviruses of subgenera B, C, D and E (Paabo *et al.,* 1986) and functions to modulate the cell-surface expression of major histocompatibility complex class I antigens. This, consequently, prevents the cell from being lysed by the adenovirus specific cytotoxic T lymphocytes (Andersson *et al.,* 1985; Burgert & Kvist, 1985; Rawle *et al.,* 1989).

ORF 10 has the potential to encode a polypeptide of approximately 37.1 KDa whose deduced amino acid sequence contained 10 canonical N-glycosylation sites and a stretch of hydrophobic amino acids at the carboxyl end, characteristic of the corresponding membrane bound E3 gp19 KDa protein encoded by Ad5 (fig. 17). Although the sequences between BAV2, Ad2 and Ad5 differ, the hydrophobicity plots generated for both Ad5 gp19 KDa and BAV2 were very similar. However, the five conserved cysteines and the 20 amino acid sequence located just upstream of the transmembrane domain (which was found to be conserved in Ad3, Ad35, Ad5 and Ad2 glycoproteins; Flomenberg *et al.,* 1988), were absent from the sequences of BAV2 ORF 10.

It is not surprising that there was no homology between the known adenoviral nucleotide or protein sequences and the predicted BAV2 E3 sequences . Belak and co-workers (1983) proposed that there were two BAV2 subspecies: subspecies A as represented by strain #19 (used in this study) and subspecies B with ORT-111, a strain originally isolated from lambs. From electron microscopic heteroduplex analysis, it was revealed that these regions were well matched except at the position which fell within the E3 region of Ad2 (Belak *et al.,* 1986).

If adenoviruses diverged at the time of speciation, one would expect that this region would vary not only within a group, but among groups and especially between viruses with different hosts. This region is probably subjected to selection pressures from the immune response and ought to vary. Different adenovirus serotypes infect different tissues and different hosts. Therefore it is not surprising that all known E3 gene products are not produced by each adenovirus; in fact , some appear to be unique to one serotype or another (Signas *et al.,* 1986). For example Ad3, Ad7 and Ad35 have two additional ORFs, which are not found in Ad2 and Ad 5 (reveiwed by Gooding & Wold, 1990).

The E3 region has been shown to be non essential for viral replication *in vitro* and *in vivo* (Morin *et al.,* 1987; Ginsberg *et al.,* 1989). This has implications for the development of BAV2 as a viral vector. The maximum amount of DNA that can be packaged in the virions is approximately 105 percent of the wild-type genome (Ghosh-Choudhury *et al.,* 1986). To incorporate larger DNA segments, it is necessary to delete appropriate amounts of viral DNA. The 2.3 kbp putative E3 region in BAV2 is larger thanthe comparable region in any other non human adenovirus studied. If this region could be deleted in BAV2 without having an effect on replication or pathogenicity of the virus, the capacity for a DNA insert would increase to 4.5 kbp. However, Ginsberg *et al.* (1989) suggested that deletions in the E3 may change the virulence of the virus and thus render them unsuitable for vaccines. It is not known whether the BAV2 E3 region is dispensable for viral replication. Therefore, it is important that this region is well characterized before it is determined whether BAV2 could be a suitable deletion vector.

The E3 region is maintained in natural isolates, suggesting that it is important in the establishment of an infection in the natural host. Viruses are known to utilize their genome efficiently and do not retain nonfunctional DNA.

Therefore, it is unlikely that a region that comprises 7 percent of the viral genome in BAV2 and 9 percent in Ad5 is nonessential. Ginsberg *et al.* (1989) proposed that the human adenovirus E3 proteins moderate an adenoviral infection, lessening its severity. It is likely that the BAV2 E3 region also plays an important role in the interaction between the virus and the host.

5.3 Organization and homology of the fibre gene region

The current evidence that ORF 14 codes for the BAV2 fibre protein is based on comparisons with the fibre coding region in other adenoviruses. Structurally there is great variability in the fiber protein between serotypes, not only in fibre length but also in the number of fibre per vertex (Norrby, 1969). The fiber protein protrudes from the surface of the virion and serves to attach it to the cell surface (Phillipson *et al.*, 1968). It is the fibre protein to which many of the host's antibodies are directed (reviewed in Phillipson *et al.*, 1983). Natural variants of the fiber protein within the virion population are probably selected by the host's response to the major antigenic determinants during infection and are, therefore, expected to vary between serotypes (Wadell *et al.,* 1986; Horwitz , 1990 a).

The deduced amino acid sequence of BAV2 was analyzed according to the model proposed by Green *et al.* (1983) for the structure of the adenovirus fibre. The fibre is divided into three functional units: (a) the tail, (b) the long narrow shaft composed of a variable number of pseudo repeats and (c) the variable knob, presumed to contain the receptor binding site. Unlike other adenoviruses studied, the first 50 amino acids are not conserved in BAV2. These amino acids make up the tail region. The tail region is highly conserved among adenovirus serotypes because it is the part of the virion that interacts with

the penton base and is consequently unexposed (Wadell *et al.,* 1986). The structure of the tail has been confirmed by the determination of the polarity and through crystallization of the fibre (Deveux *et al.,* 1987; 1990). According to the region alignment proposed (fig. 19), the tail region in BAV2 is shorter and is the least conserved amongst all adenovirus serotypes studied so far. The alignment between Ad41 and BAV2 was not entirely arbitrary. The amino acid alignments of BAV2 with Ad41, Ad40, Ad2 and Ad5 using the FASTA program were compared and were found to be similar.

The basic structural unit of the shaft is made up of 15 amino acids (Green *et al.,* 1983) The number of motifs varies between serotypes: from 6 in Ad3 (Signas *et al.,* 1985) to 42 in BAV3 (Mittal *et al.,* 1992). Such a repeating motif consisting of periodically occurring prolines, glycines and hydrophobic residues in the shaft region, is preserved in the deduced amino acid sequence of BAV2 (fig. 20). The shaft region of BAV2 showed the greatest similarity to Ad41. The 22 segments of BAV2 have hydrophobic residues predominating in positions A1, A3, B1 and B3, as noted in other adenoviruses (Green *et al.,* 1983; Signas *et al.,* 1985; Chroboczek & Jacrot, 1987;Raviprakash *et al.,* 1989; Kidd *et al.,* 1990; Dragulev *et al.,* 1991). There was a proline or glycine residue in position a5 in 20 of the 22 motifs. The proline or glycine in position a5 is conserved among adenovirus serotypes, implying that these amino acids play a critical role in the fibre's structure.

The knob region was 166 amino acids long and had an overall identity of 32 percent and a similarity of 65 percent with Ad41. The most significant homology was between residues 390 and 410. According to the proposed alignment, this would mark the transition between the shaft and the knob. These 20 residues were highly conserved between BAV2 and Ad40, Ad41, Ad5 and Ad2. Whether this conservation is a structural constraint or whether this region

may be important in the interaction of the fibre with other proteins, remains to be established.

The presumed BAV2 fibre is 557 amino acids long, resembling the size of the 547, 562, and 581 amino acid fibre proteins of Ad40, Ad41, Ad2 and Ad5, respectively (Chroboczek & Jacrot, 1987; Kidd *et al.*, 1990). The size of the BAV2 fibre has not been previously reported. The size of the polypeptide appears to determine the length of the fibre (Signas *et al.*, 1985; Chroboczek & Jacrot, 1987). Therefore, the length of the BAV2 fibre is probably similar to the length of group C and group F adenoviruses. The length of the Group C fibres determined using electron microscopy is approximately 30 nm (Norrby, 1969).

A poly A sequence was not found at the end of ORF 14. However, the sequence ATTAAA was found 6 bp away from the termination sequence. This signal may function as a poly A site for the BAV2 ORF 14. Also, in addition to the tripartite leader sequences contained in most of the late mRNAs, Ad2 fibre mRNA can contain additional leader sequences (x, y and z) encoded upstream in the E3 region (Uhlen *et al.*, 1982). Raviprakash *et al.* (1989) found sequences homologous to the Ad2 leaders in MAV1. No evidence of these sequences was found in BAV2.

There were nine possible sites for N-linked glycosylation (Asn -X-Ser/Thr) found in BAV2. Since the fibre polypeptide of Ad2 has been demonstrated to contain two residues of GlcNAc linked to the polypeptide chain by an O-glycosidic bond (reviewed by Mei & Wadell, 1993), the fibre of BAV2 may also be glycosylated.

Summary

The work conducted in this study and presented in this thesis can be summarized as follows:

1. The region between 74.8 and 90.5 map units has been sequenced.

2. The putative E3 region lies between 76.0 and 83.2 map units, it is about 2300 bp long and possibly encodes 4 polypeptides larger than 6 KDa.

3. No ORFs show significant homology with any other known DNA or protein sequence.

4. BAV2 shows 81 percent nucleotide homology to the 316 bp E3 leader sequence of Ad2.

5 The characteristic TATA box and the E3 cap site of mRNA were found in BAV2.

6. ORF 1 and 14 of BAV2 were found to be homologous to protein VIII (located on the left) and the fibre protein (on the right)of the E3 region respectively.

7. The deduced amino acid sequence of ORF 1 showed an overall identity of 54, 53, 54, and 33 percent with Ad2, Ad41, Ad5 and MAV1, respectively.

8. The last 70 AA of BAV2 ORF 1 appeared to be highly conserved, showing an identity of 63 percent and a similarity of 87 percent with Ad2 pVIII.

9. ORF 14 encodes the BAV2 fibre polypeptide with a theoretical molecular weight of 59 600 Da.

10. A comparison between the deduced amino acid sequence of BAV2 and Ad41, Ad40, Ad2 and Ad5, revealed that they have an identical secondary structure consisting of a tail, shaft and knob.

11. The amino acids 37 to 390 formed a typical shaft domain of 22, 15 residue repetitive motifs.

12 . The periodicity of the prolines and the hydrophobic residues in the amino acid sequence is preserved in BAV2.

13. The first 20 amino acids of the knob are highly conserved between BAV2 and Ad40, Ad41, Ad2 and Ad5.

LITERATURE CITED:

Adrian, T., B., Best, J.C., Hierholzer and R., Wigand (1989a). Molecular epidemiology and restriction site mapping of adenovirus type 3 genome types. *J. Clin. Micobiol.* **27:** 1329-1334.

Adrian, T., B., Best, J.C., Hierholzer and R., Wigand (1989b). Molecular epidemiology and restriction site mapping of adenovirus type 7 genome types. *Arch. Virol.* **106:** 73-84.

Andersson, M., S., Paabo, T., Nilsson and P.A. Peterson. 1985. Impaired intracellular transport of class I MHC antigens as a possible means for adenoviruses to evade immune surveillance. *Cell.* **43:** 215-222.

Baber, D.J. and J.B., Condy. 1981. Isolation and characterization of bovine adenoviruses types 3,4 and 8 from free-living African buffaloes *(Syncerus caffer). Res. Vet. Sci.* **31**:69-75.

Babich, A. and J.R., Nevins. 1981. The stability of early adenovirus mRNA is controlled by the viral 72kd DNA binding protein. Cell. 36: 371-379.

Bartha, A., 1969. Proposal of subgrouping of bovine adenoviruses. *Acta Vet. Hung.* **19:** 319-321.

Beard, CAW., A.O., Ball, E.H., Wooley and K., Spindler. 1990. Transcription mapping of mouse adenovirus type 1 early region 3. *Virology* **75:** 81-90.

Belak, S., G., Berencsi, M., Rusvay, K., Lukacs and I., Nasz. 1983. DNA structure, and hemagglutination properties of bovine adenovirus type 2 strains which bypass species specificity. *Arch. Virol.* **77:** 181-194.

Belak, S. and V., Palfi. 1974a. An adenovirus isolated from sheep and its relationship to type 2 bovine adenovirus. *Arch. ges. Virusforsch.* **46:** 366-369.

Belak, S. and V., Palfi. 1974b. Pneumoenteritis of lambs caused by adenovirus . *Acta Vet. Hung.* **24:** 327-328.

Belak, S., A., Virtanen, J., Zabielski, M., Rusvai, G., Berencsi and U., Pettersson. 1986. Subtypes of bovine adenovirus type 2 exhibit major differences in region E3. *Virology.* **153:** 262-271.

Benko, M., A. Bartha, K., Mostl and F., Burki. 1989. A heteroploid permanent cell line originating from embryonic calf thyroid supporting the replication of all known bovine adenovirus serotypes. *Vet. Microbiol.* **19:** 317-324.

Benko, M., A., Bartha and G., Wadell. 1988. DNA restriction enzyme analysis of bovine adenoviruses. *Intervirol.* **29:** 346-350.

Birnboim, H.C. and J., Doly. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* **7:** 1513-1523.

Boros, B., Z., Graf, M., Benko and A., Bartha, 1985. Isolation of bovine adenovirus from fallow deer (*Dama dama*). *Acta Vet. Hung.* **33:** 119-123.

Breathnach, R. and P., Chambon. 1991. Organization and expression of eukaryotic split genes coding for proteins. *Annual Review of Biochemistry.* **50:** 349-383.

Burgert, H-G., J.L., Maryanski and S., Kvist. 1987. "E3/19K" protein of adenovirus type 2 inhibits lysis of cytolysis T lymphocytes by blocking cell surface expression of histocompatibility class I antigens. *Proc. Natl. Acad. Sci. USA.* **84:** 1356-1360.

Burgert, H.-G. and S., Kvist. 1985. An adenovirus type 2 glycoprotein blocks cell surface expression of human histocompatibility class I antigens. *Cell* **41:** 987-997.

Burki, F., H., Wege, C., Reich- Rohrwig and B., Hinaidy. 1978. Bovine adenoviruses I. Characterization and serological classification as type 4 of two isolates from latently infected calf testicles. *Zbl. Vet. Med. B.* **25:** 555-565.

Carlin, C.R., A.E., Tollefson, H.A., Brady, B.L., Hoffman, and W.S.M., Wold 1989 Epidermal growth factor receptor is down-regulated by a 10,400 MW protein encoded by the E3 region of adenovirus. *Cell.* **57:** 135-144.

Challberg, M.D., and T.J., Kelly, Jr. 1979. Adenovirus DNA replication *in vitro.* *Proc. Natl. Acad. Sci. USA.* **76(2):** 655-659.

Chow, L.T., T.R., Broker and J.B., Lewis. 1979. Complex splicing patterns of RNAs from the early regions of adenovirus-2. *J. Mol. Biol.* **134:** 265-303.

Chroboczek, J., F., Bieber and B., Jacrot. 1992. The sequence of genome of adenovirus type 5 and its comparison with the genome of adenovirus type 2. *Virology.* **186:** 280-285.

Chroboczek, J. and B. Jacrot, 1987. The sequence of adenovirus fiber: similarities and differences between serotypes 2 and 5. *Virology.* **161:** 549-554.

Cladaras, C., and W.S.M., Wold. 1985. DNA sequence in the early E3 transcription unit of adenovirus **5:** *Virology.* **140:** 28-43.

Cladaras, C., B., Bhat, and W.S.M. Wold. 1985. Mapping the 5' ends, 3' ends, and splice sites of mRNAs from the early E3 transcription unit of adenovirus **5:** *Virology.* **140:** 44-54.

Darbyshire, J.H., 1966 a. Oncogenicity of bovine adenovirus type 3 in hamsters. *Nature* (London). **311:** 102.

Darbyshire, J.H., A.R., Jennings, P.S., Dawson, P.H., Lamont and A.R., Omar. 1966 b. The pathogenenesis and pathology of infection in calves with a strain of bovine adenovirus type 3. *Res Vet. Sci.* **7:** 81-93. .

Devaux, C., M., Adrian, C., Berthet-Colominas, S., Cusack and B., Jacrot. 1990. Structure of adenovirus fibre 1. Analysis of crystals of fibre from adenovirus serotypes 2 and 5 by electron microscopy and X-ray crystallography. *J. Mol. Biol.* **215:** 567-588.

Devaux, C., M.L., Caillet-Boudin, B., Jacrot and P., Boulanger. 1987. Crystallation, enzymatic cleavage, and the polarity of the adenovirus type 2 fiber. *Virology.* **161:** 121-128.

Dorsett, P.H. and H.S. Ginsberg. 1975. Characterization of type 5 adenovirus fiber protein. *J. of Virol.* **15(1):** 208-216.

Dragulev, B.P., S., Sira, M.G. Abouhaidar and J.B., Campbell. 1991. Sequence analysis of putative E3 and fiber genomic regions of two strains of canine adenovirus type 1. *Virology.* **183:** 298-305.

Duerksen- Hughes, P.J., T.W., Hermiston, W.S.M., Wold and L.R., Gooding. 1991. The amino-terminal portion of the CD1 of the adenovirus E1A protein is required to induce susceptibility to tumour necrosis factor cytolysis in adenovirus - infected mouse cells. *J.Virol.* **65,** 1236-1244.

Duerksen- Hughes, P.J., W.S.M., Wold and L.R., Gooding. 1989 Adenovirus E1A renders infected cells sensitive to cytolysis by tumor necrosis factor *J. Immunol.* **143,** 4193-4200

Enders, J.F., J.A., Bell, J.H., Dingle, T., Francis Jr., M.R., Hilleman, R.J., Huebner, and A.M.-M., Payne. 1956. "Adenoviruses": group name proposed for the new respiratory-tract viruses. *Science.* 124: 119-120.

Feinberg, A.P. and B., Vogelstein. 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132:** 6.

Feinberg, A.P. and B., Vogelstein. 1984. Addendum: A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **137:** 266.

Fickett, J.W, 1982. Recognition of protein coding regions in DNA sequences. *Nucl. Acids Res.* **10:** 5303-5318.

Flomenberg, P.R., M., Chen and M.S., Horwitz. 1988. Sequence and genetic organization of adenovirus type 35 early region 3. *J. Virol.* **62:** 4431-4437.

Fox, J.P., C.E., Hall and M.K., Cooney. 1977. The Seattle virus watch. VII Observations of adenovirus infections. *Am. J. Epidemiol.* **105:** 362-386.

Ghosh-Choudhury, G., Y., Haj-Ahmad, P., Brinkley, J., Rudy, and F.L., Graham. 1986. Human adenovirus cloning vectors based on infectious bacterial plasmids. *Gene.* **50:** 161-171.

Ginsberg, H.S., U., Lundholm-Beauchasmp, R.L., Horswood, B., Pernis, W.S.M., Wold, R.M., Chanock, and G.A., Prince. 1989. Role of early region 3 (E3) in pathogenesis of adenovirus disease. *Proc. Natl. Acad. Sci.* USA. **86:** 3823-3827.

Glutzman, Y. 1984. Introduction: Evolution of eukaryotic viral vectors. In Eukaryotic Viral Vectors. (Y. Glutzman, Ed.), pp. 1-5. Cold Spring Harbor: New York.

Gooding, L.R., L., Aquino, P.J., Duerksen-Hughes, D., Day, T.M., Horton, S., Yei and W.S.M., Wold. 1991a. The E1B 19,000-molecular-weight protein of group C adenoviruses prevents tumor necrosis factor cytolysis of human cells but not of mouse cells. *J. Virol.* **65(6):** 3083-3094.

Gooding, L.R., L.W., Elmore, A.E., Tollefson, H.A., Brady and W.S.M., Wold. 1988. A 14,700 MW protein from the E3 region of adenovirus inhibits cytolysis by tumor necrosis factor. *Cell* **53:** 341-346.

Gooding, L.R., T.S., Ranheim, A.E., Tollefson, L., Aquino, P., Duerksen-Hughes, T.M., Horton and W.S.M., Wold. 1991b. The 10,400- and 14,500-dalton proteins encoded by region E3 of adenovirus function together to protect many but not all mouse cell lines against lysis by tumor necrosis factor. *J. Virol.* **65(8):** 4114-4123.

Gooding, L.R. and W.S.M., Wold. 1990. Molecular mechanisms by which adenoviruses counteract antiviral defenses. *CRC Crit, Rev. Immunol.* **10:** 53-71.

Graham, F.L. and L., Prevec. 1991. Manipulation of adenvirus vectors. In Methods in Molecular Biology,Vol. 7. (Ed., E.J. Murray) pp 109-128. The Humana Press Inc., Clifton, NJ.

Green, N.M., N.G., Wrigley, W.C., Russell, S.R., Martin and A.D., McLachlan. 1983. Evidence for a repeating cross-β sheet structure in the adenovirus fibre. *EMBO* **2(1)P:** 1357-1365.

Haj-Ahmad, Y. 1986. PHD thesis. Development and Characterization of a Human Adenovirus Type 5 Cloning Vector. McMaster University, Hamilton, Ontario, Canada.

Haj-Ahmad, Y. and F.L., Graham. 1986. Development of a helper-independent human adenovirus vector and its use in the transfer of the herpes simplex virus thymidine kinase gene. *J. Virol.* **57(1):** 267-274.

Hawkins, L.K. and WSM., Wold. 1992. A 12,500 MW Protein coded by Region E3 of Adenovirus. *Virology.* **188:** 486-494.

Herisse, J., G.,Courtois and F., Galibert. 1980. Nucleotide sequence of the *Eco*RI D fragment of aenovirus 2 genome. *Nucl. Acids Res.* **8:** 2173-2192.

Herisse, J., and F., Galibert. 1981. Nucleotide sequence of the *Eco*RI fragment of adenovirus 2 genome. *Nucl. Acids Res.* **9:** 1229-1240.

Herisse, J., M., Rigolet, S., Dupont de Dinechin, and F., Galibert. 1981. Nucleotide sequence of adenovirus 2 DNA fragment encoding for the carboxylic region of the fiber protein and the entire E4 region. *Nucl. Acids Res.* **9:** 4023-4042.

Hirt, B. 1967. Selective extraction of poloma DNA From infected mouse cell culture cultures. *J. Mol. Biol.* **26:** 365-369.

Hoffman, P., P., Rajakumar, B., Hoffman, R., Heuertz, W.S.M., Wold and C.R., Carlin. 1992. Evidence for intercellular down-regulation of the epidermal growth factor (EGF) receptor during adenovirus infection by an EGF-independent mechanism. *J. Virol.* **66:** 197-203.

Hong, J.S., K.G., Mullis and J.A., Engler. 1988. Characterization of the early region 3 and fiber genes of Ad7. *Virology.* **167:** 545-553.

Horne, R.W., S., Bonner, A.A., Waterson and P., Wildly. 1959. The icosahedral form of an adenovirus. *J. Mol. Biol.* **1:** 84-86.

Horner, G.W., R., Hunter, A., Bartha, and M., Benko. 1989. A new subgroup 2 bovine adenovirus proposed as the prototype strain 10. *Acta Virol.* **109:** 121-124.

Horton, T.M., T.S., Ranheim, L., Aquino, D.I., Kusher, S.K., Saha, C.F., Ware, W.S.M., Wold and L.R. Gooding. 1991. Adenovirus E3 14.7K protein functions in the absence of other adenovirus proteins to protect transfected cells from tumour necrosis factor cytolysis. *J. Virol.* **65:** 2629-2639.

Horton, T.M., A.E., Tollefson, W.S.M., Wold and L.R., Gooding. 1990. A protein serologically and functionally related to the group C E3 14,700-kilodalton protein is found in multiple adenovirus serotypes. *J. Virol.* **64(3):** 1250-1255.

Horwitz, M.S. 1990a. Adenoviridae and their replication. In: Virology, Second Edition (B.N. Fields, D.M. Knipe *et al.,* eds) pp 1679-1721, Raven Press, Ltd., New York.

Horwitz, M.S. 1990b. Adenoviruses. In: Virology, Second Edition, (B.N. Fields, D.M. Knipe *et al.,* eds) pp. 1723-1740, Raven Press, Ltd., New York.

Hu, S.-L., W.W., Hays and D.E., Potts. 1984. Sequence homology between bovine and human adenoviruses. *J. Virol.* **49(2):** 604-608.

Igarashi, K., R., Sasada, T., Kurokawa, Y., Niiyama, K., Tsukamoto and T., Sugino. 1978. Biochemical studies of bovine adenovirus type 3 IV. Transformation by viral DNA and DNA fragment. *J. Virol.* **28:** 219-226.

Ishibashi, M. and Yasue, H. 1984 Adenoviruses of animals. In Adenoviruses, (H.S. Ginsberg ed.) Plenum Press, New York.

Key, D.W., and J.B., Derbyshire. 1984. Serological studies of parainfluenza type 3 virus, bovine adenovirus type 3 and bovine respiratory syncytial virus infection in beef calves. *Vet. Microbiol.* **9:** 587-592.

Kidd. A.H., J., Chroboczek, S., Cusack and R.W.H., Ruigrok. 1993. Adenovirus type 40 virions contain two distinct fibers. *Virology.* **192:** 73-84.

Kidd, A.H. and M.J., Eramus. 1989. Sequence characterization of adenovirus 40 fiber gene. *Virology.* **173:** 134-144.

Kidd, A.H., M.J., Erasmus and C.T., Tiemessen. 1990. Fiber sequence heterogeneity in subgroup F adenoviruses. *Virology.* **179:** 139-150.

Klein, M., E., Early and J., Zellat. 1959. Isolation from cattle of a virus related to human adenovirus. *Proc. Soc. Exp. Bio. Med.* **102(1):** 1-4.

Klein, M., J., Zellat and T.C., Michaelson. 1960. A new bovine adenovirus related to human adenovirus. *Proc. Soc. Exp. Bio. Med.* **165:** 340-342.

Krajcsi P., A.E., Tollefson, C.W., Anderson A.R., Stewart, C.R., Carlin and W.S.M., Wold. 1992a. The E3-10.4K protein of adenovirus is an integral membrane protein that is partially cleaved between Ala_{22} and Ala_{23} and has a C_{cyt} orientation. *Virology* **187:** 131-144.

Krajcsi, P., A.E., Tollefson, C.W., Anderson and W.S.M., Wold. 1992b. The adenovirus E3 14.5-Kilodalton protein, which is required for down-regulation of the epidermal growth factor receptor and prevention of tumor necrosis factor cytolysis , is an integral membrane protein oriented with its C Terminus in the cytoplasm. *J. Virol.* **66(3):** 1665-1673.

Krajcsi, P., and W.S.M., Wold. 1992c. The adenovirus E3-14.5K protein which is required for prevention of TNF cytolysis and for down-regulation of the EGF receptor contains phosphoserine. *Virology.* **187:** 482-498.

Kyte, J. and R.F., Doolittle. 1992. A simple model for displaying the hydrophobic character of a prorein. *J. Mol. Biol.* **157:** 105-132.

Lehmkuhl, H.D. and P.M., Gough. 1977. Investigation of causative agents of bovine respiratory tract disease in a beef cow-calf herd with an early weaning program. *Am. J. Vet. Res.* **38(11):** 1717-1720.

Lehmkuhl, H.D., M.H., Smith and R.E., Dierks. 1975. A bovine adenovirus type 3: Isolation, characterization and experimental infection in calves. *Arch. Virol.* **48:** 39-46.

Linne, T. 1992. Differences in the E3 regions of the canine adenovirus type 1 and type 2. *Virus Res.* **23:** 119-133.

Mamhoudi, M., and Lin, V.K., 1989. Comparison of two different hybridization systems in Northern transfer analysis. *Biotech.* **7:** 331-333.

Mattson, D.E. 1973. Adenovirus infection in cattle. *JAVMA* **163:** 894-896.

Mei, Y. and G., Wadell., 1993. Hemagglutination properties and nucleotide sequence analysis of adenovirus genome types 11p and 11a.

Mei, Y. and G., Wadell 1992. The nucleotide sequence of adenovirus type 11 early region: Comparison of genome type Ad11p and Ad11a. *Virology.* **191:** 125-133.

Mittal, S.K., L., Prevec, L.A., Babiuk and F.L., Graham. 1992. Sequence analysis of bovine adenovirus type 3 early region 3 and fibre protein genes. *J. Gen. Virol.* **73:** 3295-3300.

Mohanty, S.B. 1971. Comparative study of bovine adenoviruses. *Am. J. Vet. Res.* **32(12):** 1899-1905.

Morin, J.E., M.D., Lubeck, J.E., Barton, A.J., Conley, A.R., Davis, and P.P.,

Hung.1987. Recombinant adenovirus induces antibody response to hepatitis B virus surface antigen in hamsters. *Proc. Natl. Acad. Sci.* USA. **84:** 4626 -4630.

Motoi, M. and K., Ogawa. 1985. Choroid plexis papilloma and giant cell glioblastoma induced in hamsters with bovine adenovirus type 3 . *Acta Neuropathol.* **66:** 218-222.

Niiyama, Y., R., Sasada, K., Igarashi, T., Kurokawa and Y., Sugino. 1981. Biochemical studies on bovine adenovirus type 3. V. some properties of mouse cells transformed with viral DNA fragments. *Cell Structure and Function* **6:** 121-131.

Norrby, E. 1969. The structural and functional diversity of adenovirus capsid components. *J. Gen. Virol.* **5:** 221-236.

Paabo, S., T., Nilsson and P.A., Peterson. 1986. Adenoviruses of subgenera B, C, D, and E modulate cell- surface expression of major histocompatibility complex class 1 antigens. *Proc. Natl. Acad. Sci. U.S.A.* **83:** 9665-9669.

Panicali, D., S.W,. Davis, R.L., Weinberg and E., Paoletti. 1983. Construction of live vaccines by using genetically engineered poxviruses: Biological activity of recombinant vaccinia virus expressing influenza virus hemagglutinin. *Proc. Natl. Acad. Sci. USA.* **80:** 5364-5368.

Panigraphy S., K.J., McCormick and J.J., Trentin. 1977. Oncogenic and nononcogenic bovine adenoviruses and guanine-cytosine content of their DNA. *Am. J. Vet. Res.* **38:** 1151-1152.

Pearson, W.R. and D.J., Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci.*, USA. **88:** 2444-2448.

Persson, H., H., Jornvall and J., Zabielski. 1980. Multiple mRNA species for the precursor to an adenovirus-encoded glycoprotein: identification and structure of the signal sequence. *Proc. Natl. Acad. Sci.* USA **77:** 6349.

Pieniazek, N.J., S.B., Slemenda, D., Pieniazek, J. Velarde Jr., and R.B., Luftig. 1989. Sequence of human enteric adenovirus type 41 tak fiber protein gene. *Nucl. Acids Res.* **17(22):** 9474.

Philipson, L. (1983). Structure and assembly of adenoviruses. In Current Topics in Microbiology and Immunology , (W. Doerfler, Ed.). Vol109, pp 1-52.. SpringerVerlag, Berlin/Heidelberg.

Philipson, L., K., Lonberg-Holm and U., Pettersson. 1968. Virus-receptor interaction in an adenovirus system. *J. Virol.* **2(10):** 1064-1075.

Raviprakash, K. S., A., Grunhaus, M.A., El Kholy and M., Horwitz. 1989. The mouse adenovirus type 1 contains an unusual E3 region, *J. Virol.* **63:** 5455-5458.

Rawle, F.C., A.E., Tollefson, W.S.M., Wold and L.R. Gooding. 1989. Mouse anti-adenovirus cytotoxic T lymphocytes inhibition of lysis by E3 gp 19K but not E3 14K. *J. Immunol.* **143:** 2031-2037.

ondhuis, P.R. 1968. A new bovine adenovirus. *Arch Gesamte Virusforsch* **25:** 235-236.

Roos, R.P., 1989. Adenovirus. Handbook of Clini. Neur. **12:** 281-293.

Rowe, W.P., R.J., Huebner, L.K., Gilmore, R.H., Parrott and T.G., Ward. 1953. Isolaton of cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proc. Soc. Exp. Bio. Med.* **84:** 570-573.

Sanger,F., S., Nicklen, and A.R., Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad Sci.* **74:** 5463.

Sambrook, J., E.F., Fritsch and T., Maniatis 1989. Molecular cloning : A laboratory Manual (Cold Spring Harbour Laboratory , Cold Spring Harbour).

Scanziani, E., G., Sironi, M., Finazzi and M., Luini. 1989. Adenoviral fibrinous enteritis in calves. *Dtsch. tieraztl. Wschr.* **96:** 165-168.

Signas, C., Akusjarvi, G. and U., Petersson 1986. Region E3 of human adenoviruses: differences between the oncogenic adenovirus -3 and nononcogenic adenovirus -2 *Gene* **50:** 173-184.

Signas, C., Akusjarvi, G. and U., Petersson 1985. Adenovirus 3 fiber polypeptide gene: implications for the structure of the fiber protein. *J .Virol.* **53(2):** 672-678.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98:** 503

Stouten, P.F.W., C., Sander, R.W.H., Ruigrok and S., Cusack. 1992. New triple -helical model for the shaft of the adenovirus fibre. *J. Mol. Biol.* **226:** 1073-1084.

Tollefson, A.E., and W.S.M., Wold., 1988. Identification and gene mapping of a 14,700-molecular-weight protein encoded by region E3 of group C adenoviruses. *J. Virol.* **62:** 33-39.

Tollefson, A.E., P., Krajcsi, M.H., Pursley, L.R., Gooding and W.S.M., Wold. 1990a. A 14,500 mw protein is coded by region E3 of group C human adenoviruses. *Virology* **175:** 19-29.

Tollefson, A.E., P., Krajcsi, S., Yei, C.R. Carlin and W.S.M., Wold. 1990b. A 10,400-molecular-weight membrane protein is coded by region E3 of adenovirus. *J. Virol.* **64(2):** 794-801.

Tollefson, A.E., Scaria, A., Saha, S.K., and W.S.M., Wold. 1992. The 11,600-M_W protein encoded by region E3 of adenovirus is expressed early but is greatly amplified at late stages of infection. *J. Virol.* **66:** 3633-3642.

Tollefson, A.E., A.R., Stewart, S., Yei, S.K., Saha and W.S.M., Wold. 1991. The 10,400- and 14,500-dalton proteins encoded by region E3 of adenovirus form a complex and function together to down-regulate the epidermal growth factor receptor. *J. Virol.* **65(6):** 3095-3105.

Tooze, J. 1980. Molecular biology of tumor viruses, In DNA Tumor Viruses, second edition part 2. Cold spring Harbouir Laboratory, Cold Spring Harbour, N.Y.

Uhlen, M., C., Svensson, S., Josephson, P., Alestrom, J.B., Chattapadhyaya, U,. Pettersson and L., Phillipson. 1982. Leader arrangement in the adenovirus fiber mRNA. *EMBO J.* **1:** 249-254.

Venkatesh, L.K., M.Q., Arens, T., Subramanian and G., Chinnadurai. 1990. Selective induction of toxicity to human cells expressing human immunodeficiency virus type 1 Tat by a conditionally cytotoxic adenovirus vector. *Proc. Natl. Acad. Sci. USA.* **87:** 1746-8750.

Wadell, G., A., Allard, M., Evander, L., Quan-gen. 1986. Genetic variability and evolution of adenoviruses. *Chem. Scripta.* **26B:** 325-335.

Wang, E.W., M.O., Scott and R.P., Ricciardi. 1988. An adenovirus mRNA which encodes a 14,700-m_r protein that maps to the last open reading frame of region E3 is expressed during infection. *J. Virol.* **62(4):** 1456-1459.

Weber, J. M., B.G., Talbot and L., Delorme. 1989. The orientation of the adenovirus fiber and its anchor domain identified through molecular mimicry. *Virology.* **168:** 180-182.

Wilson-Rawls, J., S.K., Saha, P., Krajcsi, A.E., Tollefson, L.R., Gooding, and W.S.M., Wold. 1990. A 6700 mw membrane protein is encoded by region E3 of adenovirus type 2. *Virology.* **178:** 204-212.

Wold, W.S.M., C., Cladaras, S.C., Magie, and N., Yacoub. 1984. Mapping a new gene that encodes an 11,600-molecular weight protein in the E3 transcription unit of adenovirus 2. *J. Virol.* 52: 307-313.

Wold W.S.M. and L.R., Gooding. 1991. Minireview: Region E3 of adenovirus : A casette of genes involved in host immunosurveillance and virus-cell interactions. *Virology.* **184:** 1-8.

Ye, W.W., B.B., Mason, M., Chengalvala, S.-M., Cheng, G., Zandle, M.D., Lubeck, S.-G., Lee, S., Mizutani, A.R., Davis and P.P., Hung. 1991. Co-expression of hepatitis B virus antigens by a non-defective adenovirus vaccine vector. *Arch. Virol.* **118:** 11-27.

Zhang, X., A.J.D., Bellett, R., Tha Hla, A.W., Braithwaite and A., Mullbacher. 1991. Adenovirus type 5 E3 gene products interfere with the expression of the cytolytic T cell immunodominant E1a antigen. *Virology.* **180:** 199-206.

Zock, C. and W., Doerfler. 1990. A mitigator sequence in the downstream region of the major late promoter of adenovirus type 12 DNA. *EMBO.* **9(5):** 1615-1623.